Optimal Taxation with On-the-Job Search*

Jesper Bagger — Royal Holloway
Espen R. Moen — BI Norwegian Business School
Rune M. Vejlin — Aarhus University

February 15, 2018

Abstract

We provide a theoretical and empirical study of the optimal taxation of labor income in the presence of search frictions and unobservable amenities using comprehensive Danish matched employer-employee data. Heterogeneous workers undertake costly search off- and on-the-job in order to locate more productive jobs that pay higher wages. More productive workers search harder, resulting in equilibrium sorting where low-type workers are overrepresented in low-wage jobs while high-type workers are overrepresented in high-wage jobs. Absent taxes, worker search effort is efficient, because the social and private gains from search coincide. The optimal tax system balance efficiency and equity concerns at the margin. Equity concerns make it desirable to levy low taxes on (or indeed, subsidize) low-wage jobs including unemployment, and levy high taxes on high-wage jobs. Efficiency concerns limit how much taxes an optimal tax system levy on high-paid jobs, as high taxes distort the workers' incentives to search. Using detailed micro data on wages, labor market transitions, and income tax filings we estimate the structural model and compare the the actual Danish tax regime with the optimal one. Preliminary results suggest the optimal tax schedule exhibits less progressivity than the actual tax system in place. The model allows us to quantify the welfare gains from adopting an optimal income tax schedule.

Key words: Optimal taxation, On-the-Job Search, Sorting, Job ladder, Matched Employer-Employee Data.

JEL codes: H21, H31 J63, J64.

1 Introduction

The equilibrium allocation of resources does not materialize costlessly in a market with search frictions, as it does in a perfectly competitive market. That is, improving the

*We would like to thank Vidar Christiansen for comments on an early draft.

allocation of resources in a frictional market requires resources in itself, and is therefore value creation. This value creation may be distorted by taxation. In the context of a frictional labor market, workers and firms undertake costly search in order to identify the jobs in which their productivity (and therefore, wage) is the highest. Taxes on labor income will have an allocative effect, as it reduces the incentives of workers to search for more high-productive, well-paid jobs.

We study the optimal taxation of labor income in a frictional labor market where workers search to find a better job both while employed and unemployed.. Workers are *ex ante* heterogeneous in their productivity, and hence face different job prospects. A benevolent planner has a preference for income equality, where income is measured as (expected) lifetime discounted income. On the one hand, in order to redistribute from high- to low-productive workers, the planner wants to levy high taxes on individuals with a high current income (which tend to be high-productive workers) and redistribute the tax revenues towards agents with a low current income (which tend to be low-productive workers).

On the other hand, workers of all types expend resources to find more productive jobs that pay better. All workers enter as unemployed, and through job search, first as unemployed and then as employed, gradually climb the wage/productivity ladder. Hence, search effort yields both private and social returns, but is costly and—crucially—not deductible (hence on-the job search may well be undertaken during leasure time). Increasing taxes on wages at a given rung of the wage/productivity ladder reduces the incentives to search at lower rungs, but does not affect the incentives to search on higher rungs. Hence, in order to protect the creation of value, the planner wants to levy lower taxes on higher wage levels. The optimal tax system trades off the equity and efficiency concerns of the planner.[1]

We first construct a model in which both unemployed and employed workers search (on-the-job search). In this model, search is random, however in the absence of taxes the equilibrium is constrained efficient. Workers of different types face different wage distributions when they search. The role of firms are very much played down, they do not contribute actively to the search process, and wages are equal to productivity. We first derive analytically conditions for an optimal tax system. We then extend the model by allowing for amenities, which are not observable by the planner and hence

[1]We assume that taxation does not distiort other margins, like the intensive and the extensive margin of labor supply. Hence our focus is solely on the workers' search decisions. In a richer model with endogenous labor supply, the distortionary effects of taxes on the workers' search decisions will come in addition to (the well-studied) distortions in labor supply.

cannot be taxed. When choosing between jobs, a worker takes both wages net of taxes and amenities into account, and this creates a new margin that is distorted by taxes. Equity concerns imply that the planner levy taxes that increase in gross wages. After successful search, a worker therefore has a tendency to accept too few jobs that offer a higher wage and too many jobs that offer a lower wage than the wage in the current job, relative to what the planner would like the worker to do. This is because workers, in contrast to the planner, consider wages net of taxes when choosing between jobs. In isolation this effect tends to increase the efficiency loss of taxes. Again we derive analytical conditions for an optimal tax system. We solve the basic model (without amenities) numerically, and give examples of optimal tax systems.

Finally, the model is estimated on Danish matched employer-employee data set with detailed information on individual level income tax rates. The data cover the entire Danish population of workers and firms during 1999-2003, a period where the income tax regime in Denmark is stable, and contain information on daily individual labor market histories, hourly wages, and—crucially—detailed information on individual tax filings. This part of the data allow us to simulate individual-level marginal tax rates (as in Kleven and Schultz, 2014). Hence, we estimate the structural model using the actual tax function faced by individual workers, including individual circumstances such as marital status, non-labor income, and individual level deductions.[2] We construct an Indirect Inference where we fit a number of moments related to workers' labor market behavior. Identification of the structural parameters is discussed.

With the structural parameter estimates in hand we proceed to compute the optimal tax regime for the estimated economy, and compare this to the actual 1999-2003 Danish tax regime. Preliminary results suggest the optimal tax schedule exhibits less progressivity than the actual tax system in place. The model allows us to quantify the welfare gains from adopting an optimal income tax schedule.

The existing literature on optimal taxation and search is mostly concerned with the search decisions of unemployed workers. Hungerbuhler, Lehmann, Parmentier, and van der Linden (2006) analyse optimal taxation in a one-shot unemployment search model. In their model, firms use resources to open vacancies and wages are determined by wage bargaining. They assume (like us) that workers are risk neutral, while the planner has preferences over the (expected) income distribution over different worker types. They show how a revelation mechanism can be applied at the bargaining stage, so that the worker and the firm bargain over which type to reveal to the planner. As

[2]We are grateful to Henrik Kleven and Esben Schultz for making their tax simulator available to us.

a result, the revelation principle can be used to derive the optimal mechanism. Under the optimal taxation scheme, the employment level is optimal for the most productive worker-firm pairs, while there is over-employment for the lower types that do search.

Golosov, Maziero, and Menzio (2013) study optimal taxation in a one-shot competitive search equilibrium model with identical, risk averse workers and heterogeneous firms. There is a fixed cost for workers from sending an application. The equilibrium without taxation is inefficient, as optimal risk sharing requires that workers are compensated for applying to jobs they do not get. In the constrained efficient equilibrium, the unemployment insurance is set so that workers are indifferent between searching for any job and not searching, as this gives maximum insurance given workers' incentive compatibility constraint. There is no transfers between workers searching for different firm types; the firms in effect finance the unemployment benefit of the workers they attract but do not hire. As a result, taxes are regressive.

Shi and Wen (1999) analyze the effect of taxes in a model of random unemployment search, in which workers accumulate capital. Higher labor taxes discourage working, and leads to lower investments by firms and lower wages. Capital taxation on the other hand increases labor supply, as workers get a lower return on their capital. Hence capital taxation may improve the allocation of resources. Domeji (2005) analyzes optimal taxation within the same modeling framework, and find that the optimal capital tax is zero if and only if the Hoisio's condition is satisfied. Jiang (2012) uses a similar setup to analyse the welfare effects of a UK tax reform. Arseneau and Chugh (2012) studies taxation in a calibrated DSGE model with search frictions, and argues that cyclical variations in the search-based labor wedge call for taxes that vary over the business cycle. Wilemme (2017) studies taxation in a model of mismatch, and shows that taxes should be regressive to correct for workers not being sufficiently selective. Geromichalos (2015) study optimal taxation with risk averse workers in a one-shot urn ball model of the labor market.

A couple of recent papers analyze taxation and on-the-job search. Sleet and Yazici (2017) studies optimal taxation in a Burdett and Mortensen (1998) model of on-the job search. A tax on labor income reduces net wages for workers, and hence increases their pre-tax reservation wage. As a result, the entire wage distribution shifts, and this influences the division of rents between workers and employers. Bagger, Hejlesen, Sumiya, and Vejlin (2017) evaluates equilibrium effects on labor allocation of a series of tax reforms in Denmark and also analyze optimal tax reforms using an equilibrium on-the-job search model with Burdett and Mortensen (1998) wage setting.

On a conceptual level, our paper is also related to papers outside the search literature. Saez (2002) analyses a model of taxation in which taxes influence participation (the extensive margin) as well as which firm type (level) to work for (the intensive margin). Working for a firm at a higher level gives higher income, but this may come at a cost. If the extensive margin is sufficiently important, taxes for low-income employed workers may be lower than for unemployed workers. This is studied in more detail in Christiansen (2015). Although the Saez (2002) model is very different from ours, there are interesting similarities between the two: In Saez' model, reducing taxes at a given level induces some workers who previously were choosing an occupation one level above or below switches to that level. In our model, by contrast, reducing taxes at a given job type reduces the incentives to search for workers in that job type, increases the incentives to search for all workers further down the in the job ladder, and leaves the incentives unchanged for all workers higher up in the job ladder.

Another related paper is Best and Kleven (2013). They study an environment with learning by doing, so that the wage of an old worker depends on his labor supply as young. A higher marginal income tax on old workers thereby reduces the labor supply of young workers. The model is calibrated to wage data for young and old workers from the PSID. Estimates of the effect of labor supply on future wages are obtained from existing studies of the effect of experience on wages. The authors find that with age-dependent taxes, the presence of on-the-job learning implies that an optimal tax system prescribes lower taxes for old workers than for young workers. With age-independent taxes, it implies that the optimal tax system prescribes a lower marginal tax rates across the board.

2 Basic Model

There are J worker types in the economy. The fraction of workers of type j is denoted by τ_j, i.e. $\sum_{j=1}^{J} \tau_j = 1$. Workers discount the future at a pure discount rate $\bar{r}$ and exit the market at rate λ to be replaced by new unemployed workers. The effective discount rate is thus $r \equiv \bar{r} + \lambda$. Workers search equally efficient off- and on-the-job. The number of firms is exogenously given.

The job ladder has $n+1$ rungs with rungs indexed by $i \in \{0, 1..., n\}$. Each rung i is associated with a productivity level $y_i > y_{i-1}$, with the lowest rung $i = 0$ being the value of home production (i.e. productivity during unemployment). After successful search, the probability that a worker of type j draws a job of productivity y_i is denoted f_i^j, and define $F_i^j \equiv \sum_{k=1}^{i} f_k^j$. Our assumption that productivities are discretely distributed is made for tractability only, extending the analysis to allow for a continuum of job productivities is straight-forward. The offer distribution is independent of the search intensity and acceptance decisions of workers, and is a primitive of the model.

Our modeling of the productivity distribution of workers and firms is general, and may capture different assumptions regarding the productivity distribution of firms, complementarities between worker and firm types, and of match-specific productivity components. Let us give an example of an underlying structure that gives rise to a job distribution of our type: Suppose there are L firm-types in the economy. The fraction of firms of type ℓ is ξ_ℓ. All firms exhibit constant returns to scale in production. The productivity of a type-j worker and a type-ℓ firm is stochastic and revealed at the point when they meet. Denote by $\tilde{f}(j,\ell) \geq 0$ the probability that the realized productivity is y_i. It follows that $f_i^j = \sum_{l=1}^{L} \xi_l \tilde{f}(j,l)$. This production structure, without any restrictions on the probabilities, are sufficiently rich to capture the production functions of most on-the-job search models in the literature.

Worker types are indexed so that a higher j means a higher type. More precisely we assume that if F^{j+1} stochastically dominates F^j for all j. In addition we assume that f_i^{j+1}/f_i^j is strictly increasing in i whenever properly defined ($f_i^{j'} \neq 0$) and different from zero.

An employed worker may be hit by a negative employment shock and enter unemployment. The rate at which this happens is s. In principle, s may depend on both worker- and job type, in which case we write $s = s_i^j$.

The arrival rate of job offers, p, to a worker is proportional to the number of efficiency units of search, e, provided by the worker. That is, $p = Ae$, where A is an exogenous constant. We may think of A as proportional to the number of firms K in the economy, $A = aK$. The aggregate number of matches is $NA\bar{e}$, where N is the aggregate number of workers in the economy (which we will normalize to 1), and $\bar{e}$ their average search intensity. Hence our search technology differs from the standard search technology in that a worker's search intensity does not create congestion effects and thereby lower job finding rates for other workers. This seems a natural assumption if workers search for firms, and firms have a constant-return-to-scale production technology and give a job offer to all workers that approach them. The cost of search effort is $\kappa_j c(e)$, where the coefficient κ_j may be worker-type dependent. If not, we normalize κ to 1.

The wage a worker receives gross of taxes is equal to his productivity y_i. This is consistent with a model of wage bargaining in which the worker has all the bargaining power. It follows that absent taxes, there are no externalities associated with the workers' choice of search intensity. Higher search intensity of a worker does not create congestion externalities for other workers, and although it increases the hiring probability of firms, this does not create positive externalities as the worker's wage is equal to his productivity.[3]

We assume that the government cannot observe worker types directly and restrict the tax system to be contingent on current wages. Hence the government, through taxation, determines the net-of-tax wage w_i that a worker (irrespective of his or het type) receives in a job of productivity y_i. Note that w_0 is the net-of-tax unemployment income. We require that w_i is nondecreasing in i, so that income taxation leaves the ranking of job unaltered. It follows that the probability distribution of w_i is the same as that of y_i.[4]

An important issue is the preferences of the agents. As in Golosov, Maziero, and Menzio (2013), Best and Kleven (2013) and others we assume that agents in the economy are risk neutral, while the planner's welfare weights on the different groups of workers depend on their expected net present income.[5] Let W_0^j be the net present income of an unemployed worker of type j less of search costs. Note that W_0^j is also the lifetime income of a worker that enters the market. An entering type-j worker's objective is to maximize W_0^j. A planner's objective is to maximize the welfare function

$$\Omega = \sum_{j=1}^{J} \kappa_j \Phi(W_0^j) \tag{1}$$

where $\Phi(W_0^j)$ is strictly increasing and concave.

A driving assumption in our analysis is that search effort is unobservable to the planner, and cannot be contracted upon. Furthermore, we assume that search cost is associated with distress when searching for a job and reduced leisure, not reduced income. Hence the search cost is born fully by the worker, while the gain in terms of

[3]The search-and wage determination processes are special cases of those developed by Diamond (1982), Mortensen (1982), and Pissarides (1985), with an exogenous number of vacancies and with the workers' bargaining power equal to 1. See also Pissarides (1994) for a model with on-the-job search.

[4]A simpler way to proceed is to impose a particular functional form of the tax function, for instance by following Heathcote, Storesletten, and Violante (2017) and assume that the tax function is given by $z - \lambda z^{1-\tau}$. It follows that the after-tax-income of a worker is $\lambda z^{1-\tau}$ and the elasticity of the after-tax-income is $1 - \tau$.

[5]One way of justifying this assumption is that idiosyncratic risk is shared in families of workers, where all the family members are of the same type.

higher income is partly taxed away. We also assume that the search cost is independent of the worker's income. Hence the level of wages in the different jobs do not influence search intensity, only the wage differentials. At this point we follow Saez (2002) when he assumes that the choice of sectors do not depend on the income levels in the different sectors, only the differences in income between them.

3 Optimal taxation

In this section we will first solve the model for a given set of taxes, and then derive the optimal tax system.

3.1 Asset values

For any variable X_i^j, define $EX_i^j = \sum_{k=i+1}^n f_k^j X_k^j$ and the operator Δ as $\Delta X_i^j = \sum_{k=i+1}^n f_k^j (X_k^j - X_i^j)$ for any variable X^j. Let W_i^j denote the expected discounted income of a type-j worker in a type-i job. Analogously, let M_i^j denote the expected discounted future income gross of taxes for a type-j worker in a type-i job. It follows that

$$
\begin{aligned}
(r+s)W_i^j &= w_i + Ae_i^j \Delta W_i^j + sW_0^j - c(e_i^j) && (2)\\
c'(e_i^j) &= A\Delta W_i^j && (3)\\
(r+s)M_i^j &= y_i + Ae_i^j \Delta^j M_i^j + sM_0^j - c(e_i^j) && (4)
\end{aligned}
$$

For a given vector $(w_0, ...w_n)$, and a given W_0^j, the model can be solved recursively for each type separately:

1. At the top:

$$
\begin{aligned}
W_n^j &= \frac{w_n + sW_0^j}{r+s}\\
e_n^j &= 0\\
M_n^j &= \frac{y_n + sM_0^j}{r+s}
\end{aligned}
$$

2. Recursively further down

$$\begin{aligned} W_i^j &= \frac{w_i + Ae_i^j EW_i^j + sW_0^j - c(e_i^j)}{r+s+Ae_i^j(1-F_i^j)} \\ c'(e_i^j) &= A\Delta W_i^j \\ M_i^j &= \frac{y_i + Ae_i^j EM_i^j + sM_0^j - c(e_i^j)}{r+s+Ae_i^j(1-F_i^j)} \end{aligned}$$

Lemma 1 *For any vector $(w_0, w_1, ...w_n)$ the vector $W_0^j, ...W_n^j$ exists and is unique.*

Proof. For a given $\hat{W}_0^j$, equation (2) -(4) uniquely define $W_i^j = W_i^j(\hat{W}_0^j)$ for all i. In particular, $W_0^j = W_0^j(\hat{W}_0^j)$. Equilibrium is a fixed-point to this mapping. As the mapping is continuous and defined on a compact and convex set, Brouwer's fixedpoint theorem ensures existence. Furthermore, Maxwell's sufficiency condition is satsified $(W_i^j(\hat{W}_0^j + \Delta) < W_i^j(\hat{W}_0^j) + k\Delta)$ for some $k < 1$), which implies that the fixed-point is unique. ■

Consider now a change in w_i. For a worker in a job $k > i$, this only influences the npv income W_0^j, and since the transition rate to unemployment is s independently of the job type, it follows that

$$\frac{dW_k^j}{dw_i} = \frac{s}{r+s}\frac{dW_0^j}{dw_i} \quad \forall k > i. \tag{5}$$

From the envelope theorem it follows that the effect of an increase in w_i on W_i^j is

$$\frac{dW_i^j}{dw_i} = \frac{1 + s\frac{dW_0^j}{dw_i}}{r+s+Ae_i^j(1-F_i^j))}. \tag{6}$$

The effect of a change in w_i on W_{i-l} is recursively defined as

$$\frac{dW_{i-l}}{dw_i} = \frac{Ae_{i-l}^j + s\frac{dW_0^j}{dw_i}}{r+s+Ae_{i-l}^j(1-F_{i-l})} \sum_{k=i-l+1}^{i} f_i^j \frac{dW_{i-k}^j}{dw_i} \tag{7}$$

An increase in w_i reduces search effort at w_i, and increases search effort at lower wages. From (3) it follows that

$$\frac{de_{i-l}^j}{dw_i} = A\frac{d\Delta W_{i-l}}{dw_i}/c''(e_{i-l}^j) \tag{8}$$

Remark 1 $\frac{d\Delta W_{i-l}}{dw_i} > 0$ *for all* $l \geq 1$.

Proof. We know that this is true for $l = 1$. Suppose it is true for all $l < l'$. Suppose it is not true for l'. Then we know from (2) that $\Delta W_{l'}^j$ decreases, but that can only be true if $W_{l'}^j$ increases, a contradiction. ■

Remark 2 *The higher is the worker type, the higher is the search intensity at a given current wage level* w_i.

Proof. Consider worker types j and $j+1$. Since F_{j+1} stochastically dominates F_j. Hence for all i, $W_i^j < W_i^{j+1}$. From equation (2) it follows that $\Delta W_i^{j+1} > \Delta W_i^{j+1}$. ■

Remark 3 *Note that* ΔW_i^j *is independent of* W_0^j. *Define* $\tilde{W}_i^j$ *by (2), but with* 0 *substituted in for* W_0^j. *It follows that* $W_i^j = \tilde{W}_i^j + \frac{s}{r+s} W_0^j$. *This simplifies the derivatives, as it is sufficient to calculate the derivative of* $\tilde{W}_i^j$ *and then add the derivative of* W_0^j *at the end. In particular it follows that we can write* $rW_0^j = w_0 + Ae_i^j(\tilde{W}_i^j + \frac{s}{r+s} W_0^j)$. *This simplifies the calculation of the derivative of* $\tilde{W}_i^j$.

Finally, consider M_i^j. Note that a change in w_i does not affect M_i^j directly, only through its effect on e_i^j.[6] From (4) we have that

$$\begin{aligned} (r+s)M_i^j &= y_i + e_i^j A \Delta M_i^j + sM_0^j - c(e_i^j) \\ &= y_i + e_i^j A \Delta W_i^j - c(e_i^j) + e_i^j A \Delta (M_i^j - W_i^j) + sM_0^j \end{aligned} \tag{9}$$

From the envelope theorem it follows that $\frac{dM_i^j}{dw_i}$ is given by

$$(r+s)\frac{dM_i^j}{dw_i} = \frac{de_i^j}{dw_i} A \Delta (M_i - W_i) + s\frac{dM_0^j}{dw_i} \tag{10}$$

For $k > i$, it follows that

$$\frac{dM_k}{dw_i} = \frac{s}{r+s}\frac{dM_0}{dw_i} \tag{11}$$

[6] Recall that, due to the envelope theorem, marginal changes in e_i^j has no effects on W_i^j. Note the similarity with the literature on sufficient statistics for welfare analysis, see Chetty (2009).

Further down the job ladder, it follows that

$$(r+s)\frac{dM_{i-l}}{dw_i} = A_{i-l}e^j_{i-l}\sum_{k=i-l+1}^{i}\frac{dM_{i-k}}{dw_i}f^j_i + \frac{de^j_{i-l}}{dw_i}\Delta(M^j_{i-l}-W^j_{i-l}) + s\frac{dM^j_0}{dw_i} \quad (12)$$

A change in w_i does not directly influence M^j_i, only indirectly through e^j_i. A change in e^j_i influences bot the return to search and the cost of search. Taking the derivative of (9), using the first order condition for the worker's choice of e^j_i (and include the effect of unemployment) gives (10).

The effects of a change in w_i on M_{i-l} are first that effects of changes in the values $M^j_{i-l+1}, ...M^j_i$ this is reflected in the first term. Second, a change in w_i also influences the search effort, which is captured by the second term. The last term captures the effects through M^j_0.

Finally, note that the equivalent of Remark 3 applies for M^j_i.

Remark 4 . *Suppose taxes* $(y_i - w_i)$ *are constant (in dollar) at and above* w_i *and that* $s = 0$. *Then* $\Delta(M^j_i - \Delta W^j_i) = 0$, *and hence* $\frac{dM^j_i}{dw_i} = 0$. *If taxes are proportional, then* $\frac{dM_i}{dw_i} < 0$.

3.2 The planner's problem

Suppose the planner in steady state needs to raise an amount $\bar{M}$ in net present value (NPV) income.[7] The planner takes gross incomes $(y_0, ..., y_n)$ as given, and chooses a vector $(w_0, ..., w_n)$ of net wages so as to maximize welfare. The planner's problem therefore reads

$$\begin{aligned}
\max_{w_0,...,w_n} \quad & \sum_{j=1}^{n}\tau_j\Phi(W^j_0) \\
& \text{S.T.} \\
\sum_{j=1}^{n}\tau_j M^j_0 - \bar{M} &\geq \sum_{j=1}^{n}\tau_j W^j_0 \\
c'(e^j_i) &= A\Delta W^h_i \\
w_i &\leq w_{i+1}
\end{aligned}$$

[7] Equivalently, suppose the planner needs to raise a steady state flow revenue of $r\bar{M}$.

The first constraint says that the planner cannot use more than the total value of resources available. Note that as $c(e)$ enters linearly in both W and M, it cancels out. Hence the constraint requires that the income available must be equal to the total income generated. In what follows we assume that wages are increasing in types, and then check that they actually do in optimum afterwards. The Lagrangian reads (ignoring the constraint $w_i \leq w_{i+1}$)

$$L = \sum_{j=1}^{n} \tau_j \Phi(W_0^j) + \lambda \left(\sum_{j=1}^{n} \tau_j M_0^j - \sum_{j=1}^{n} \tau_j W_0^j - \bar{M} \right) - \sum_{i,j} \mu_i^j (A \Delta W_i^j - e_i^j), \tag{13}$$

where λ and μ_i^j are the Lagrangian parameters associated with the constraints. The first order condition for w_i reads (where $\frac{dM_0^j}{dw_i}$ is given by 12, and thus includes the effects of changes in effort levels)

$$\sum_{j=1}^{n} \tau_j (\Phi'(W_0^j) - \lambda) \frac{dW_0^j}{dw_i} = -\lambda \sum_{j=1}^{n} \tau_j \frac{dM_0^j}{dw_i}. \tag{14}$$

The left-hand side is the welfare effect of increasing w_i over and above the shadow value of income. The right-hand side is the cost in terms of reduced search effort.

Intuitively, the value of $\frac{dW_0^j}{dw_i}$ depends on how large fraction of the time (discount rate weighted) that a worker of type j spends in a job of type i. The lower is the worker type, the more time the worker spends in unemployment, as inflow to unemployment typically is independent of or decreasing in a worker' type j type while outflow is higher for high-type workers with a higher search effort. This bias may be even stronger in low-type jobs. Low-type workers have a higher probability of meeting a low-type job, and hence may have a higher inflow rate to these jobs than high-type workers. At the same time the outflow rate is still higher for the high-type workers.

Let $g_j = \Phi'(W_0^j)$, $\dot{W}_i = \sum_{j=1}^{n} \tau_j \frac{dW_0^j}{dw_i}$, and $\dot{M}_i = \sum_{j=1}^{n} \tau_j \frac{dM_0^j}{dw_i}$, and finally let $\omega_i^j = \frac{W_0^j}{dw_i} / \dot{W}_i$. Then we can rewrite (14) as

$$\sum_{j=1}^{n} \tau_j (g_j - \lambda) \omega_i^j = -\lambda \frac{\dot{M}_i}{\dot{W}_i} \tag{15}$$

From the planner's perspective, increasing w_i has two effects: a distributional effect and an incentive effect. The left-hand side of (15) reflects the distributional effect, it shows how an NPV-dollar used on increasing w_i is distributed on the different worker types. A high correlation between the welfare weights g_i and the distribution weights ω_i^j implies a positive distributional effect of increasing w_i. The right-hand side shows the efficiency loss of increasing w_i, i.e., the effect on the total amount of available resources per npv dollar spent on increasing w_i^j. As we have seen, the efficiency loss tends to be positive for low values of i and negative for high values of i. The efficiency concerns thus puts a limit on how much the planner wants to redistribute.

At the lower end of the wage hierarchy, low-type workers are overrepresented. Hence for lower wages, ω_i^j tend to be decreasing in worker type. Hence the correlation between g_j and ω_i^j tends to be positive, and this calls for low (or negative) taxes at low wage levels. On the other hand, high wages /low taxes at the lower end of the wage distribution tends to reduce search effort, while higher wages at high wage levels tend to increase search effort. Efficiency considerations therefore set a limit on how much the planner wants to redistribute.

The unemployment state is somewhat special, as all agents start as unemployed, and all "restart" as unemployed after a negative employment shock. Hence it may well be that the fraction of high-types among the unemployed is higher than the fraction among workers in the low-paid jobs. In addition, increasing the income of unemployed workers clearly has a negative incentive effects for unemployed workers. Hence it is not *a priori* clear that the planner will set the unemployment benefit particularly high, but rather obtain redistribution by subsidizing low-wage jobs. This will be explored numerically below.

Remark 5 *Suppose the planner has linear preferences, so that the welfare weights ω_i^j are equal for all j independently of W_i^j. Then the planner will set $y_j - w_j = \bar{M}/r$ for all j. With such a tax policy, the planner will not distort the workers' search effort, and efficiency is obtained. This is equal to a poll tax.*

4 Including amenities

Suppose now that a job comes with two attributes, productivity y_i and other qualities, which we denote by z and refer to as amenities. Amenities cannot be observed by the planner, and hence cannot be taxed. The utility flow of a job is thus the sum of wages net of taxes and amenities. The joint distribution of amenities and productivity types for a person of type j can be written as $F^j(i, z)$. Let F_i^j denote the marginal probability that a worker of type j draws a job of productivity type (or just type) i or lower. Let f_i^j denote the probability that $y = y_i$. Let $F^j(z|i)$ denote the conditional distribution of z given i. In order to simplify the analysis we assume that the draws of productivities and amenities are independent. This means that we can write the distribution of z as $G(z)$. We will modify this assumption in future work.

Above, we required that that taxes will not influence the workers' ranking of jobs. This is consistent with the assumption that marginal taxes are less than 100 percent. However, with amenities the tax system will influence the ranking of jobs, and lead to new inefficiencies; in the presence of high marginal taxes, workers may turn down job offers with a high pre-tax wage that scores low on amenities that the planner wants the workers to accept. We still require that taxes do not alter the ordering of the pecuniary returns of jobs.

The utility flow of the current job is equal to the sum $w_i + z$. The worker will accept a new job if and only if the characteristics of the new job (i', z') satisfies $w_{i'} + z' > w_i + z$. Hence $w_i + z$ is a sufficient statistics for the npv value of the job.

Define $\tilde{z}_{i,k}(z) = w_i + z - w_k$. A worker in a job with characteristics (w_i, z) accepts a new job offer at wage level w_k if and only if the level of amenities z' in the new job satisfies $z' > \tilde{z}_{i,k}(z)$. The NPV income of a worker of type j in a job with characteristics (w_i, z) can thus be written as

$$(r + s)W_i^j(z) = w_i + z + Ae_i^j \Delta W_i^j(z) + sW_0^j - c(e_i^j) \tag{16}$$

where

$$\Delta W_i^j(z) = \sum_{k=1}^{n} f_k^j \int_{\tilde{z}_{i,k}(z)} (W_k^j(\tilde{z}) - W_i^j(z)))g(\tilde{z})d\tilde{z} \tag{17}$$

As above we have that

$$c'(e) = A\Delta W_i^j(z) \tag{18}$$

Let $\hat{F}_i^j(z)$ denote the probability that a worker of type j in a job at level i with amenities

z does not accept a job offer when it arrives. It follows that $\hat{F}_i^j(z) = \sum_{k=1}^n f_k^j G(\tilde{z}_{i,k}(z))$. From the envelope it follows that

$$W_i^j\prime(z) = \frac{1}{r + s + Ae^j(1 - \hat{F}_i^j(z))} \tag{19}$$

Lemma 2 *for a given distribution F_i^j with finite support, the equilibrium exists.*

Proof. Let us sketch the proof. The value functions $W_i^j(z)$ are real functions defined on a closed subset of R^{n+1}. Let D denote the set of continuous functions defined on that subset and bounded above by $(y_n + z^{max})/r$. We know that D is complete under the sup norm. Equation (16)-(18) define a mapping $\Gamma : D \to D$. It follows trivially that Γ is continuous, and increasing. Let $\bar{e} < \infty$ denote the supremum of e. Since the set of value functions is bounded it follows that $\bar{e}$ is bounded. We want to show that Blackwell's sufficient condition holds. Let δ be a strictly positive constant. For any $W \in D$ it follows that $\Gamma(W + \delta) \leq \Gamma W + \frac{r}{r+A\bar{e}}\delta = W + \beta\delta$ with $\beta < 1$. Hence the Blackwell sufficient condtion is satisfied, and Γ is a contraction mapping. From the contraction mapping theorem it follows that Γ has a unique fix-point. ■

Let us consider the effect on $W_i^j(z)$ of a change in w_i. Without amenities, this would only influence the NPV values in firms with wage at or below w_i (in addition to the effects through W_0). With amenities this is no longer the case, as workers may accept jobs with lower wages. As the envelope theorem still applies, the effect on effort level and acceptance decisions do not have first order effects on W^j and can be ignored. Hence

$$(r+s)\frac{dW_i^j(z)}{dw_i} = 1 + \sum_{k=1}^n f_k^j \int_{\tilde{z}_{i,k}(z)} \left(\frac{dW_k^j(\tilde{z})}{dw_i} - \frac{dW_i^j(z))}{dw_i}\right) g(\tilde{z})d\tilde{z} + s\frac{dW_0^j}{dw_i}$$

The first term on the rhs is the direct effect of an increase in the wage. The second term is the indirect effect, showing the effect on asset values after a job switch, which is influenced by a change in w_i as the worker may switch back to a job at level w_i in the future. Note that $\frac{dW_i^j}{dw_i} \neq W_i^{j\prime}(z)$, as the latter does not take into account that the value of staying in a job of type i will be higher if the worker is rehired at a job on this level in the future.

Remark 6 *In the limit, as the number of job types go to infinity and $f_k^j \to 0 \ \forall\, k$, the probability that a worker returns to a job of type i if accepting a better job offer* *goes to zero. In this case* $\frac{dW_i^j}{dw_i} = W_i^{j\prime}(z)$ *given by* (19)*. By defining* e *as a function of* $w+z$*, taking the derivative of* e *and inserting for* $\frac{dW_i^j}{dw_i}$ *gives a second order differential equation in* e *that can be solved, at least numerically.*

The effect of an increase in w_i on workers hired in other job types read

$$(r+s)\frac{dW_l^j(z)}{dw_i} = \sum_{k=1}^{n} f_k^j \int_{\tilde{z}_{i,k}(z)} (\frac{dW_k^j(\tilde{z})}{dw_i} - \frac{dW_l^j((z))}{dw_i})g(\tilde{z})d\tilde{z} + s\frac{dW_0^j}{dw_i}$$

The expression is more complicated than one may first expect, and this is due to the endogenous search effort, which varies between states. As en example, suppose i is a high-type job and $w_l + z$ is low. The search effort initially is is high, and hence an increase in w_i has a relatively high impact on $W_i^j(z)$. If the worker transfers to a job of type $k \neq i$ with a higher $w_k + z$, the search effort falls, and the effect of an increase in w_i is smaller. This feeds back into $\frac{dW_l^j(z)}{dw_i}$. Note also that in this case, we cannot abstract from the third term when the number of types increase. True, the probability of moving to state i goes to zero, but that is also true for $\frac{dW_l^j(z)}{dw_i}$. Relatively speaking, the third term does not vanish.

The total income a worker earns gross of taxes can be written as

$$rM_i^j(z) = y_i + z + Ae_i^j(z)\Delta^j M^j(z+y_i) - c(e_i^j(z)) \tag{20}$$

where $\Delta^j M_i^j(z)$ is defined analogous with $\Delta^j W_i(z)$, i.e.,

$$\Delta M_i^j(z) = \sum_{k=1}^{n} f_k^j \int_{\tilde{z}_{i,k}(z)} (M_k^j(\tilde{z}) - M_i^j(z)))g(\tilde{z})d\tilde{z}$$

Note that the threshold $\tilde{z}_{i,k}(z)$ is determined so as to maximize W^j, not M^j.

Consider an increase in w_i. Above this influenced M_i^j through its effect on search effort of workers at ring i and below. Now it may potentially influence search intensity in jobs above ring i as well. In addition we get a new effect through the workers' ranking of alternatives. Taxes drive a wedge between productivity and wage, and this may distort the worker's job acceptance decision. If taxes (in levels) are increasing with the wage – as optimal taxes typically prescribe – a worker will put less weight on gross wages (i.e. productivity) and more on amenities than the planner will. Hence a worker will have a tendency to accept too few jobs that offer a higher wage than his current wage, and too many jobs that offer a lower wage.

Formally, the derivative of (20) reads (after adding and subtracting W as above), see equation (9)

$$\begin{aligned}(r+s)\frac{dM_i^j(z)}{dw_i} &= \frac{de_i^j(z)}{dw_i}A\Delta(M_i - W_i) + Ae_i^j(z)\frac{\sum_{k=1}^n f_k g(\tilde{z}_{i,k}(z))(t_i - t_k)}{r+s+Ae_i^j(z)(1-\hat{F}_i^j(z))} \\ &+ Ae_i^j(z)\sum_{k=1}^n f_k^j \int_{\tilde{z}_{i,k}(z)} (\frac{dM_k^j(\tilde{z})}{dw_i} - \frac{dM_i^j(z))}{dw_i})g(\tilde{z})d\tilde{z} \\ &+ s\frac{dM_0^j}{dw_i} \end{aligned} \tag{21}$$

The first third and fourth terms are as above, the new term is the second term. It shows the effect of a wage change on the acceptance decision of job offers. An increase in w_i decreases the acceptance rate of other job offers, and this will give a gain/loss that depends on the difference in tax rate. This difference can be written as $\sum_{k=1}^n f_k g(\tilde{z}_{i,k}(z))(M_k^j(\tilde{z}_{i,k}(z)) - M_i^j(z))$. Note that $(r+s)M_i^j(z)) = w_i + t_i + \Delta^j M^j(z + y_i) - c(e_i^j(z))$. Since we are looking for the marginal switch, the NPV income and search and acceptance behaviour is the same before and after the switch. It follows that the flow change in value created is the difference in taxes, and that the flow difference ends at rate $Ae_i^j(z)(1-\hat{F}_i^j(z))$.[8] The third term reflects the impact on the value of future jobs, which again reflects that the worker may return to a job of type i. The final term shows the effect of an increase in w_i on the value of being unemployed.

Remark 7 *In the limit, as the number of job types go to infinity and $f_k^j \to 0\ \forall\ k$, the probability that a worker returns to a job of type i if accepting a better job offer goes to zero. In this case the third and fourth terms vanish.*

Finally, consider a change in w_i on $M_k^j(z)$ the value of being in job $l \neq i$. The first thing to note is that an increase in w_i will not influence the choice between two jobs

[8] An alternative formulation of the total income is the following.: Let $T_i^j(z)$ denote the expected discounted future tax payments of a worker of type j in a job of type i with amenities z. It follows that $M_i^j(z) \equiv W_i^j(z) + T_i^j(z)$. With this formulation, the derivatives derived above follows easily.

w_l and w_k, $k, l \neq i$. It follows that

$$
\begin{aligned}
(r+s)\frac{dM_l^j(z)}{dw_i} &= \frac{de_i^l(z)}{dw_i}A\Delta(M_l - W_l) + Ae_l^j(z)\frac{f_i g(\tilde{z}_{l,i}(z))(t_i - t_l)}{r+s+Ae_i^j(z)(1-\hat{F}_l^j(z))} \\
&+ \sum_{k=1}^{n} f_k^j \int_{\tilde{z}_{i,k}(z)} (\frac{dM_k^j(\tilde{z})}{dw_i} - \frac{dM_l^j(z))}{dw_i})g(\tilde{z})d\tilde{z} \\
&+ s\frac{dM_0^j}{dw_i} \qquad (22)
\end{aligned}
$$

Finally, consider the planner's problem. Formally, the problem is equivalent to the planner's maximization problem without amenities, with the old asset values replaced with the new ones. In particular, the trade-off between equity and incentives to search is still present. However, there are new elements in the trade-off related to the workers' acceptance decision of firms.

With the possible exception of the state of being unemployed, taxes will typically be increasing in job types because of equity considerations. As a result, workers will have a tendency to accept too few high-wage job offers (with low amenities) and accept too many low-wage offers (with high amenities). This effect is particularly prevalent at the highest paid job, jobs for which unambiguously too few workers employed in other job types accept and too many workers quit from after successful on-the-job search. Hence the introduction of amenities increases the social cost of redistribution, as there is a new margin, acceptance rate of jobs, that taxes distort (in addition to search intensity).

5 Simulation of Model Without Ameneties

In this section we show a simulation of the model without amenities.

5.1 Parameterization

The unit of time is a year. Following Christensen, Lentz, Mortensen, Neumann, and Werwatz (2005) we assume that the search cost function $c(e)$ is given by

$$c(e) = \frac{c_0}{1+1/c_1}e^{1+1/c_1}$$

We assume that y_i is uniformly distributed on the interval $[y_l, y_h]$ with $y_1 = y_l$ and $y_n = y_h$. Each worker type j draw from a worker specific offer distribution, F^j. This is approximated by a beta distribution with parameters α_j and β_j.

The social welfare function is given by

$$\Gamma = \sum_{j=1}^{J} \begin{cases} \tau_j log(W_0^j) & \text{if } \gamma = 1 \\ \tau_j \frac{(W_0^j)^{1-\gamma}}{1-\gamma} & \text{if } \gamma \neq 1 \end{cases}$$

where W_0^j is the value function for a type j worker in unemployment. To make the illustration easier we set the number of worker types to 3 ($J = 3$) and the number of firm types to 10 ($n = 10$). Each type is equally likely so $\tau_j = 1/3$.

Table 1 shows the the parameter values. We set $s = 0.2$ such that the average employment length is 5 years. A is set to hit an unemployment rate of around 10 in the no-tax case. We normalize $c_0 = 1$, since this cannot be separately identified from A. Following the results in Christensen, Lentz, Mortensen, Neumann, and Werwatz (2005) we set $c_1 = 1$ such that the search cost function is quadratic. We normalize $y_l = 1$ and set the highest level of productivity to be 5 times higher.

Table 1: Parameters

Parameter	Description	Value
$s_i^j = s$	Job destruction rate	0.2
r	Effective discount rate	0.05
A	Scale parameter in search technology	0.85
$\kappa_j = \kappa$	Worker type efficiency	1
c_0	Scale parameter in search cost function	1
c_1	Elasticity of search cost function	1
y_l	Lower bound on productivity distribution	1
y_h	Upper bound on productivity distribution	5
UI_{fact}	$y_0 = UI_{fact} \times y_l$	0.5
$\bar{M}$	Amount to raise	$0.3 \times \sum_{j=1}^{J} \tau_j \sum_{i=1}^{n} g_j^i M_i^j$
γ	Constant relative risk aversion	1
α_1	Worker type 1, Offer Distr., Alpha Parameter	0.75
α_2	Worker type 2, Offer Distr., Alpha Parameter	0.85
α_3	Worker type 3, Offer Distr., Alpha Parameter	0.95
β_1	Worker type 1, Offer Distr., Beta Parameter	5
β_2	Worker type 2, Offer Distr., Beta Parameter	4
β_3	Worker type 3, Offer Distr., Beta Parameter	3

$\bar{M}$ is set such that the government set taxes to raise 30 of the output in the no-tax equilibrium.

We are going to refer to workers of type 1, 2, and 3 as *low*, *medium*, and *high*, respectively, since $F^3 \leq F^2 \leq F^1$.

5.2 Results

In this section we compare fout different equilibria; 1) one without any taxation, 2) one with a poll-tax, 3) one with proportional taxation, and 4) one with optimal taxation. In the first equalibrium $\bar{M} = 0$, since there is no taxation to finance it, while in the last three cases we set $\bar{M} > 0$ and to the same amount such that in each of the tax schemes the government collects the same revenue.

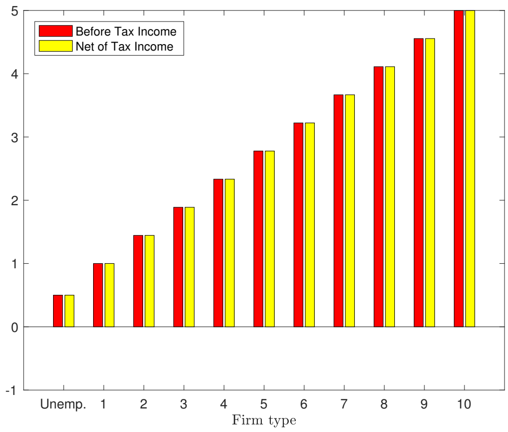
Table 2 shows some key numbers for the different equilibria. We can see that the unemployment rate differs across worker types with workers of type 1 being most

Table 2: Key Equilibrium Objects

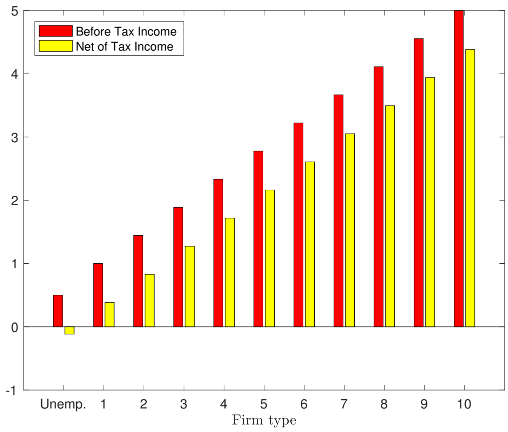
	Worker Type 1	Worker Type 2	Worker Type 3
No Tax ($\bar{M} = 0$)	-	-	-
- Unemployment Rate	0.129	0.119	0.108
- Average Before Tax Income	2.252	2.590	3.002
- Average Net of Tax Income	2.252	2.590	3.002
- W_0^j	35.134	40.513	47.457
- Ω	3.707	-	-
Poll Tax ($\bar{M} > 0$)	-	-	-
- Unemployment Rate	0.129	0.119	0.108
- Average Before Tax Income	2.252	2.590	3.002
- Average Net of Tax Income	1.636	1.974	2.387
- W_0^j	22.824	28.202	35.147
- Ω	3.342	-	-
Proportional Tax ($\bar{M} > 0$)	-	-	-
- Unemployment Rate	0.153	0.141	0.128
- Average Before Tax Income	2.135	2.460	2.864
- Average Net of Tax Income	1.576	1.816	2.114
- W_0^j	24.259	27.979	32.856
- Ω	3.337	-	-
Optimal Tax ($\bar{M} > 0$)	-	-	-
- Unemployment Rate	0.132	0.123	0.113
- Average Before Tax Income	2.221	2.544	2.934
- Average Net of Tax Income	1.647	1.925	2.242
- W_0^j	23.806	28.317	33.850
- Ω	3.345	-	-

unemployed. However, the unemployment rates do not differ that much, since all workers have similar job destruction rates. Imposing a poll-tax does not distort choices, but average net of tax income decreases and thus the value of unemployment decreases. Proportional taxation clearly distorts search choices as illustrated by the increase in unemployment and the decrease in average net of tax income. The optimal tax schedule decrease unemployment and increase net of tax income compared to the proportional tax scheme. Finally, Table 2 shows Ω, the welfare as evaluated by the planner. We note that the welfare ranking is as expected. The proportional tax regime delivers the lowest welfare, with the poll tax regime begin second. The optimal tax regime delivers the highest welfare by balancing equity and efficiency concerns at the margin.

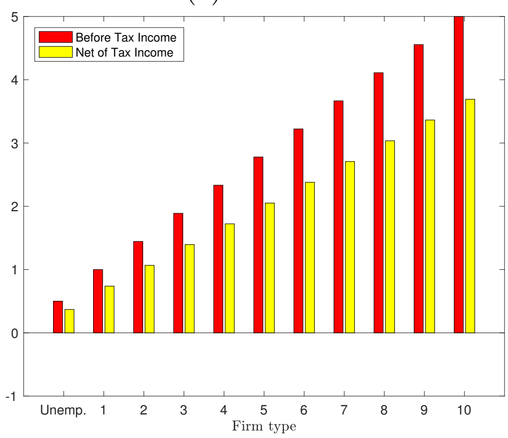
In figure 1 we show the output, which is equal to the wage income before taxation, of each state. Output is uniformly distributed across firms and we imposed a UIB level of 50 percent of the lowest output. We also show the net of tax income. In the poll-tax case all workers pay the same and unemployed workers actually receives negative payments. In the proportional tax case all workers pay 26 percent of their income. Finally, in the optimal tax case workers in all states pay taxes with unemployed worker paying y_0 in taxes. Our simulations are set up such that all worker types face (almost) the same unemployment risk as shown in Tables 1 and 2. Thus, taxing workers in this state is a relatively efficient way of collecting taxes while adding more incentives to search for higher paying jobs that are also taxed. This effect my be watered down or vanish altogether depending on the precise specification of the job destruction process.



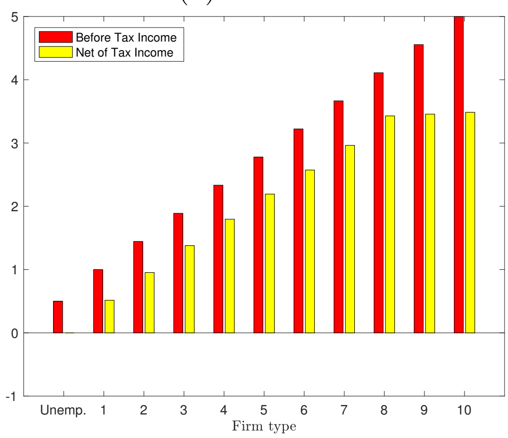
(a) No Tax



(b) Poll Tax



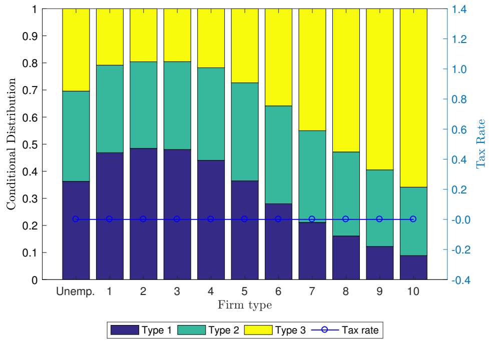
(c) Proportional Tax



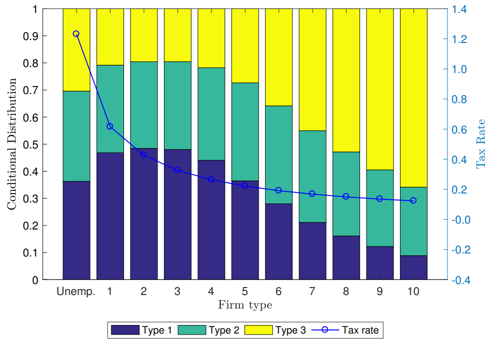
(d) Optimal Tax

Figure 1: Gross and Net Income

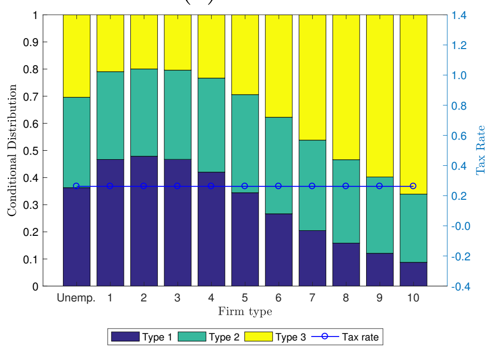
Figure 10 shows tax rate on the right side y-axis. This is just calculated as the percentage difference between the net of tax income and the before tax income as shown in figure 1. The bars show the distribution of worker types conditional on the state of the worker. The state conditional distributions are quite stable across the different tax regimes. Note, however, as we shall see further below, this does not imply taxation has no impact on the allocation of labor. However, we can conclude that the response to the different tax regimes is similar across worker-types. For the optimal tax scheme, the tax is decreasing up until the 80th percentile of firms. After the 80th percentile the tax rate start to increase. The reason is that in those firms around 50-65 percent of the workers are of the highest type. Thus, taxing workers in this state is a very good way of directing taxation at the highest type.



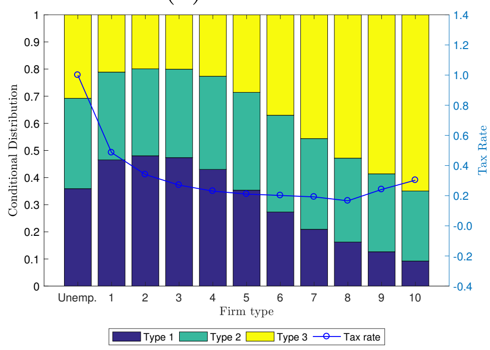
(a) No Tax



(b) Poll Tax



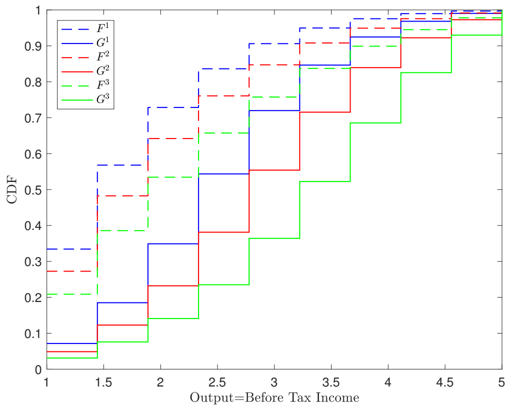
(c) Proportional Tax



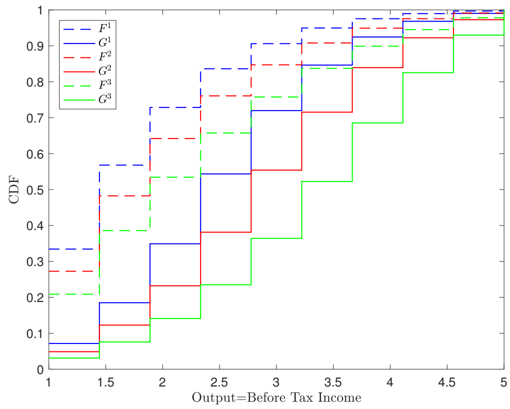
(d) Optimal Tax

Figure 2: State Conditional Worker Distribution

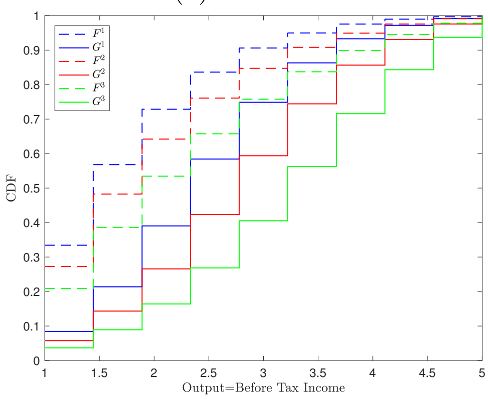
In figure 3 we show the CDF's of the offer (F) and steady state (G) distributions for each taxation scheme. The difference between the offer and steady state distributions is largest for the high worker type. This reflects that the incentives to search for a better job is highest for this type, which we show in figure 4. It is clear that taxation only influences search intensity in the proportional case, whereas it does not change search intensity that much in the optimal tax case. This is a clear indication that designing the tax scheme is important in order to minimize the cost.



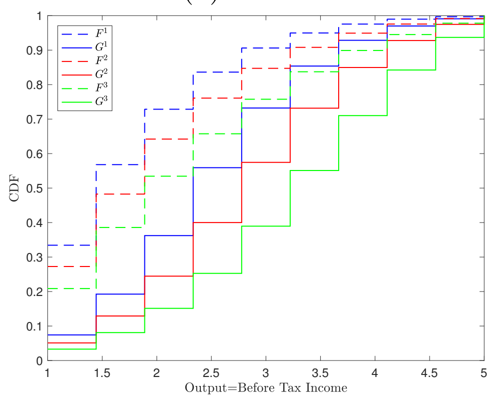
(a) No Tax



(b) Poll Tax



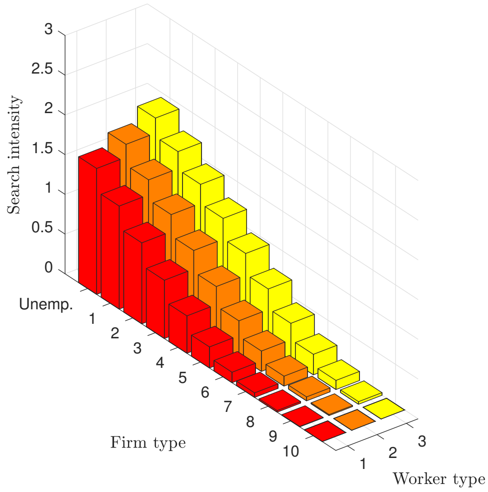
(c) Proportional Tax



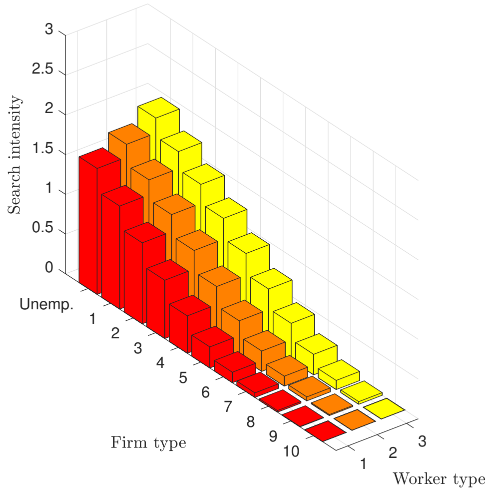
(d) Optimal Tax

Figure 3: Distribution of Workers Across Firm Types

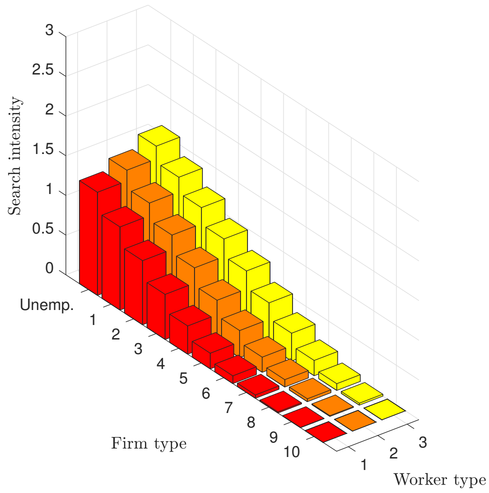
This point is further illustrated by Figure 4 which shows the search intensities for every possible worker-firm type combination (including unemployment) and for each of the four tax regimes we consider. We note that the proportional tax regime depresses search for every worker types at all rungs at the job ladder relative to the nondistorted search choices in panels (a) and (b). Under the optimal tax regime, workers search harder at the bottom and middle rungs, but less at the higher rungs, reflecting the progressivity of the optimal tax system for very high wages that allow the planner to target high-type workers.



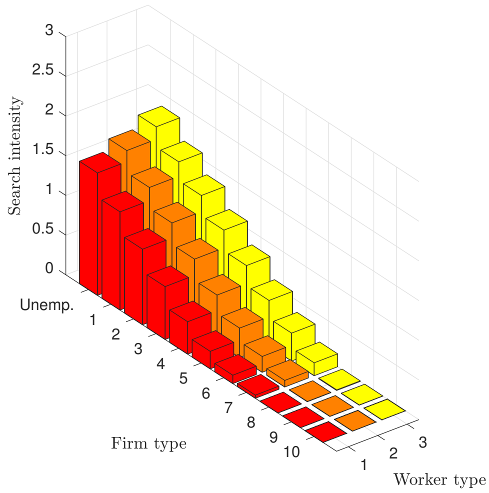
(a) No Tax



(b) Poll Tax



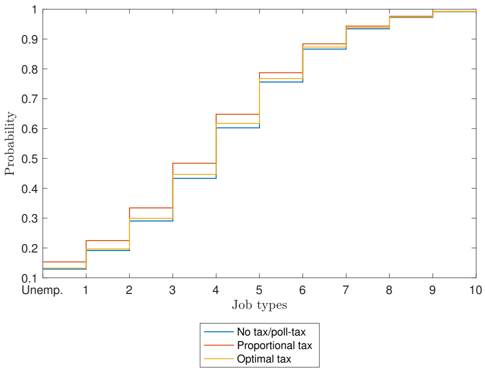
(c) Proportional Tax



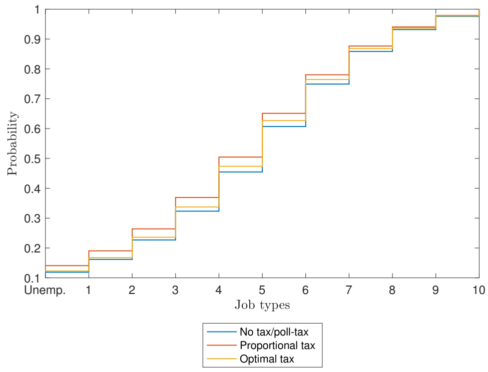
(d) Optimal Tax

Figure 4: Search Intensity

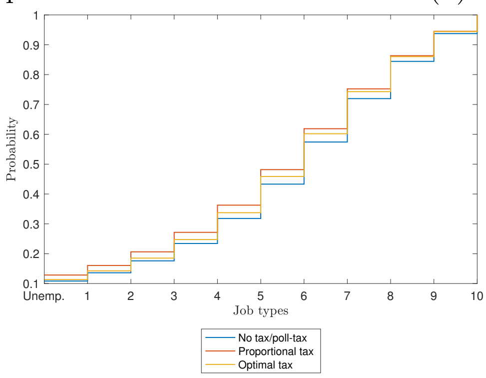
Figure 5 illustrates how different tax regimes impact the allocation of labor. Figures 5a, 5b, and 5c shows the steady state distribution of type 1 (low), type 2 (medium) and type 3 (high) workers across firm types, i.e. across the job ladder. We note that for every worker type, the labor allocation under the optimal tax regime strictly and clearly dominates the allocation under a (revenue neutral) proportional tax. The allocation under the non-distorted poll-tax regime of course dominates the allocation under the optimal tax regime.



(a) Worker type 1



(b) Worker type 2



(c) Worker type 3

Figure 5: Labor allocation

Finally, in figure 6 we show how the optimal tax schedule depends on the coefficient of relative risk-aversion in the welfare function. Recall that $\gamma = 1$ was the baseline value.

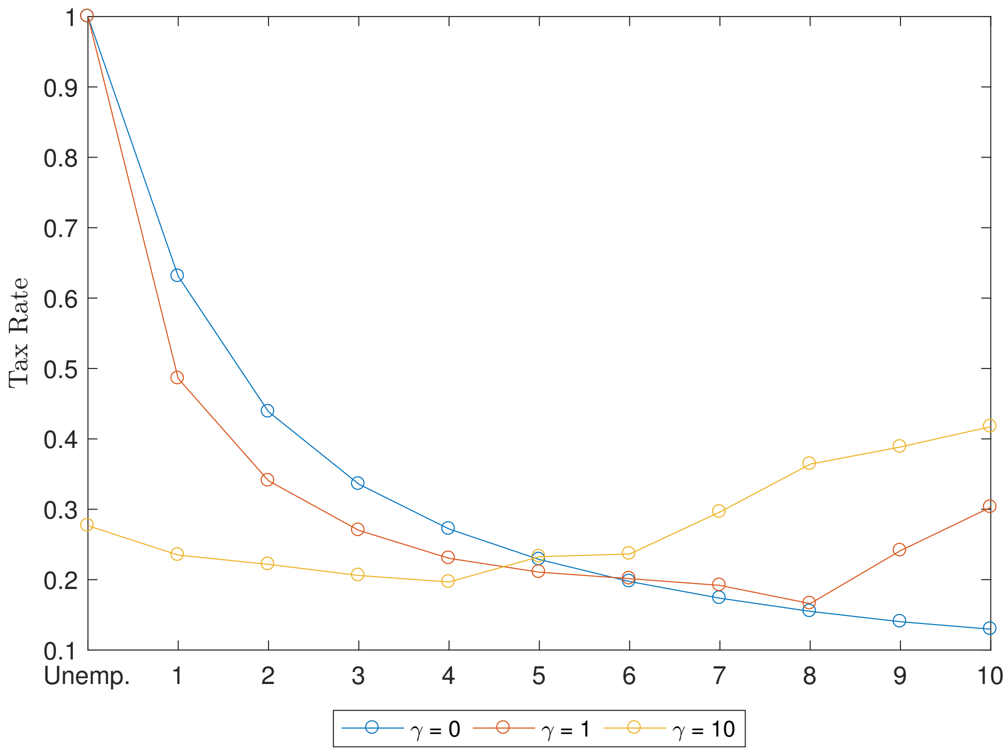
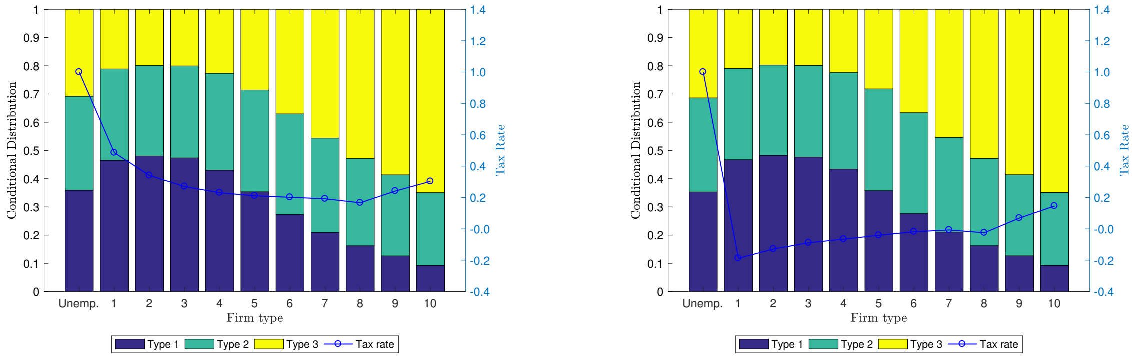


Figure 6: Optimal Tax Schemes: γ

5.3 Robustness Checks

In this section we show results for the optimal taxation if $\bar{M} = 0$ and $UI_{fact} = 0.95$.

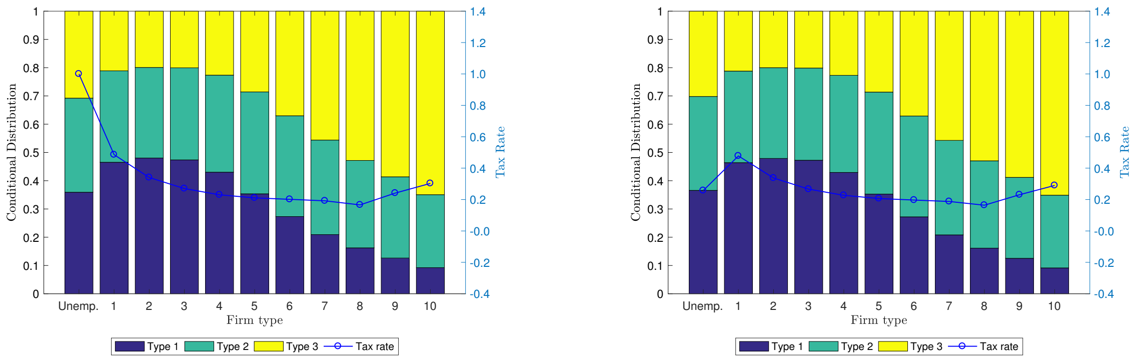
First, we set $\bar{M} = 0$ in order to see how this affects optimal policy. This corresponds to the case, where all tax-income is redistributed through transfers. The state conditional tax rate is decreasing from unemployment as it was previously, but the drop is much larger than in the baseline case with $\bar{M} > 0$. The reason is that redistributing taxes to workers in low firm types is now the most efficient way of allocating ressources to the low worker type.



(a) Optimal Tax, baseline, $\bar{M} > 0$ (b) Optimal Tax, $\bar{M} = 0$

Figure 7: State Conditional Worker Distribution

We now set $UI_{frac} = 0.95$ instead of 0.5. This does not change that much, which is only natural since all the income was taxed away for the unemployed in the baseline case. This is still optimal to do, since it is a good way of taxing the high type workers without distorting search incentives.



(a) Optimal Tax, baseline, $UI_{frac} = 0.5$ (b) Optimal Tax, $UI_{frac} = 0.95$

Figure 8: State Conditional Worker Distribution

6 Data

Our empirical analysis uses an administrative matched employer-employee dataset covering the entire Danish population during 1999-2003. We chose this period, since it has a stable income tax system. On the worker side, the dataset contains individual labor market histories measured at a daily frequency, job-specific annual average hourly wages, detailed information on individual tax filings, and a vector of background characteristics such as gender, education, income measures, and marital status. We use a tax simulator coded by Kleven and Schultz (2014) to compute individual marginal tax rates. On the firm side, the data contains information on value added at an annual frequency as well as some background characteristics, e.g. industry.

6.1 Data sources

We build our matched employer-employee panel from three sources, (i) labor market spell data, including hourly wage information, (ii) information from IDA[9], a Danish register-based matched employer-employee database maintained by Statistics Denmark containing socioeconomic information on workers and some background information on firms, and (iii) information on firms' sales and purchases from firm-level value added tax (VAT) accounts administered by the Danish tax authorities. We link the data sources using worker and firm identifiers.[10]

Labor market spell data. The labor market spell data contains individual job and non-employment spells. Information on job spells is available for the period 1985-2013 for all legal Danish residents aged 15-74, and is obtained by combining a number of administrative registers.[11] As mentioned above, for this paper, we restrict attention to the period 1999-2003. A job spell is defined as a continuous period of primary employment at a given firm.[12] Job spells, and therefore labor market transitions, are measured at the firm-level. A firm may consist of multiple workplaces. Continuous employment at different workplaces within a firm is considered as a single job spell. Nonemployment spells are periods where no job spells are recorded. We are not able to distinguish between nonemployed who actively search for a job and nonparticipants. Thus, we lump the two groups together into non-employment and impose sample selection rules described below to select those with a strong labor market attachment. The unit of observation in the labor market spell data is a person-spell-year. The job spell data includes worker and firm identifiers, start- and end-dates of the job, the annual earnings pertaining to the job, as well as an estimate of the annual number of hours worked in the job.[13] Hence, we can compute (an estimate of) the average annual wage

[9]IDA is an acronym of Integreret Database for Arbejdsmarkedsforskning.

[10]Worker identifiers are obtained from the social security registry (the CPR-registry) and firm identifiers are based on the Danish business registry (the CVR-registry).

[11]Henning Bunzel at Aarhus University has been instrumental in developing the job spell data described here. The data is described in more detail in Bunzel and Hejlesen (2016).

[12]Primary attachment is evaluated calendar month by calendar month. For each individual in each month, the primary employer is defined as the firm at which the individual works the highest number of hours in the current and next two calendar months.

[13]Annual hours are estimated using information on mandatory pension contributions. Lund and Vejlin (2015) develop and implement a procedure for computing annual hours in a job in the IDA data for the

rate in each job.

We overwrite short non-employment spells which occur between a job to job transition between two different firms or between a recall, i.e. between two consecutive job spells at the same firm. For job to job transitions we overwrite up to 14 days of non-employment and for recalls we overwrite up to 12 weeks of non-employment.

IDA data. IDA consists of several sub-panels. We use the sub-panels IDA-P and IDA-S. IDA-P contains annual background information from public registers on all individuals aged 15-74 residing legally in Denmark on the 31st of December. We retain IDA-P information on age, gender, highest completed education including date of completion, and information on any ongoing education subsequent to the highest completed education. We use education information to identify number of years of schooling for each individual, which will allow us to account for observed educational heterogeneity in labor market behavior. Date of completion of the highest completed education is used to define labor market entry. The unit of observation in IDA-P is a person-year. IDA-S contains background information from public records on all physical workplaces in Denmark. Some jobs involves work that is carried out at changing locations. Statistic Denmark designate such jobs at taking place at fictitious workplaces, and the fictitious workplaces are excluded from IDA-S.[14] We retain information on industry affiliation and a public sector indicator from IDA-S. Our analysis is carried out at the firm-level. We take a firm's industry affiliation and public sector status to be the industry and public sector status of its largest workplace.[15] The unit of observation in the (aggregated) IDA-S panel is a firm-year.

VAT data. Data on firms' sales and purchases are obtains from the panels MOMS and MOMM, constructed from firm's VAT accounts maintained by the Danish tax authorities. Danish firms settle VAT accounts monthly, quarterly or annually depending on the size of their revenue. MOMS covers the period 1995-2000 and contains annual sales and purchases. MOMM is a monthly panel starting in January, 2001. We use MOMM data up until December, 2005, aggregating the monthly information to an

period 1980-2007, primarily using information on mandatory pension contributions. We adapt this procedure for the spell data with some minor simplifications.

[14]Approximately 95% of workers are working at non-fictitious workplaces, i.e. has a fixed physical workplace.

[15]The size of the workplace is the number of workers with primary employment at the workplace on the 28th of November, as measured in IDA-S.

annual frequency.[16] Putting MOMS and MOMM together, we obtain an annual panel of firms' sales and purchases for the period 1999-2003. We compute annual firm-level value added as annual sale less annual purchases. The unit of observation in the VAT panel data is a firm-year.

Merging the data sources. We briefly describe how we link the three data sources described above. The point of departure is the labor market spell data. Recall that this dataset contains both person and firm identifiers (at least for job spells). We first merge IDA-P information onto the labor market spell data by person identifier and year and retain only person-years that are found in both data sources. This is not restrictive as 99% of the labor market spell data observations (for persons in the age range 15-74) are matched with an IDA-P observation. When we henceforth refer to the labor market spell data, we refer in fact to this intersection between the original labor market spell data and the IDA-P data.

Next, we merge the labor market spell data with the IDA-S data by firm identifier and year.

Finally, we merge VAT information onto the labor market spell data, also using the firm identifier.

6.2 Selecting the analysis data

We select our analysis data from the raw matched employer-employee panel described above to get rid of observations with missing relevant information, and to obtain a population for which our (steady state) equilibrium labor market model provides a reasonable description of labor market behavior.

We start by discarding all observations on individuals never observed with either age or education information.[17] We also discard individuals with implausible education information, defined as age minus years of schooling being less than 5. We then define labor market entry to occur at the observed date of graduation from highest completed education, or at January, 1st of the year an individual turns 19, whichever occur at the latest date. All pre-entry observations are discarded. We define labor market exit to occur at the last year an individual is observed living in Denmark (i.e. present in

[16] For firms that settle VAT accounts on quarterly or annual frequencies, the monthly information in MOMM are imputed by Statistics Denmark. As all firms settle VAT accounts at least annually, the annual aggregate data is not affected by the imputation.

[17] The primary cause of missing education data is foreign educational credentials. Hence, we are likely to discard predominantly immigrants and refugees in this step.

IDAP) or at December, 31st of the year an individual turns 55, whichever occur first. All post-exit observations are discarded. We right censor observations at age 55 to avoid labor market behavior driven by retirement considerations.

We select individuals with relatively stable labor market attachment, who primarily work full time, and who primarily work in the private sector. We are able to identify public sector employers through information on ownership structure obtained from IDA-S as described above.[18] For each individual, we delete all spells for workers who, (i) in any year worked less than 25 hours a week on average, (ii) at any point in time work in the public sector, or (iii) at any point in time does not live in Denmark (i.e. not present in IDAP).

Finally, we trend value added and wages to 2003 prices using the internal deflator computed from employment-weighted wages.[19][20]

6.3 Marginal tax functions

For estimation we will need to use the empirical tax function. Here we explain how we obtain it.

Simulating individual marginal labor income tax rates. An individual's marginal tax rate depend on several factors, including the distribution of income across the income concepts, region of residence, marital status, and, if married, spousal income. Fortunately, our data contains all the information necessary to simulate the marginal tax rates faced by individuals in our data. The simulations are carried out using a simulator of the Danish tax system constructed by Kleven and Schultz (2014).[21]

The tax simulator allow us to compute the total tax liability for each individual, for each year that we consider (i.e. 1999-2003). For the purpose of this paper, where we focus on labor income taxation, it is useful to represent individual i's tax liability in year t as $T_t(LI_{it}, \mathbf{Z}_{it})$, where $T_t(\cdot)$ is the tax function in year t, LI_{it} is individual labor income, and $\mathbf{Z}_{it}$ is a vector of other income concepts relevant for the tax liabilities in

[18] We positively identify public sector firms, effectively treating firms with missing information on ownership structure as private sector firms.

[19] We prefer to use one deflator for both wages and value added instead of two separate deflators. Using the wage deflator gives raise to the possibility that average productivity is not constant over the sample period.

[20] Wage internal inflator is 0.82, 0.90, 0.91, 0.96.

[21] We are grateful to Henrik Kleven and Esben Schultz for making the tax simulator available. The tax simulator consists of a set of SAS-programs which we downloaded from the website of the *American Economic Journal: Applied Economics*. We collected data on regional taxes in Denmark for the period to use as an input in the tax simulations.

year t. Following Kleven and Schultz (2014) we compute individual marginal tax rates as

$$T_t'(LI_{it}, \mathbf{Z}_{it}) = \frac{T_t(LI_{it} + 100, \mathbf{Z}_{it}) - T_t(LI_{it}, \mathbf{Z}_{it})}{100}.$$

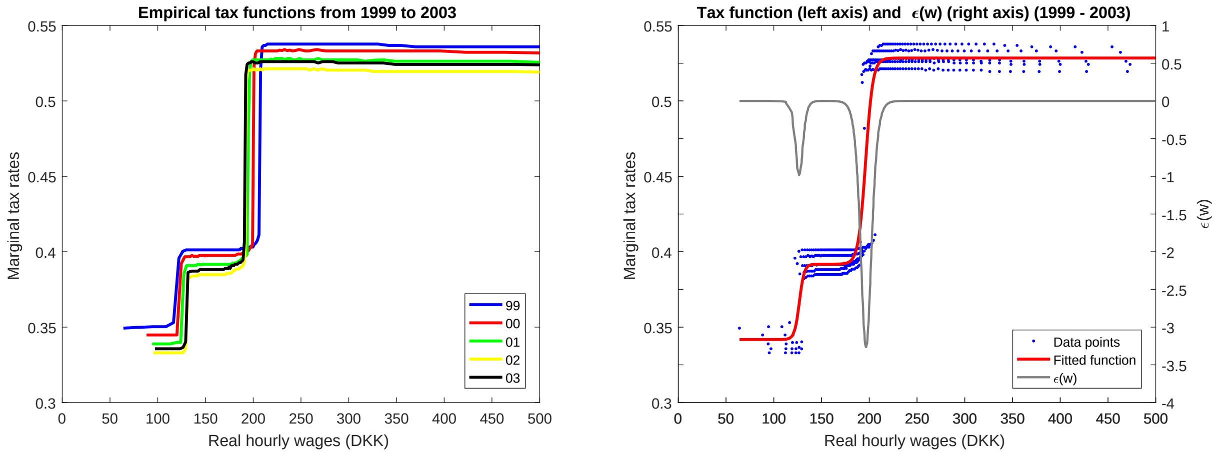
That is, we add DKK100 (approximately USD15 in 2003) to annual labor income LI, and record the share of this marginal increase that is appropriated by taxes. We stress that marginal tax rates are computed individual by individual, year by year, and reflects individual circumstances such as the composition of income across the income groups (e.g. capital income), region of residence, and relevant household characteristics. Hence, the simulated marginal tax rates represent the actual marginal tax rate faced by each individual in our data.

Estimating the marginal labor income tax function. We use the (heterogeneous) annual individual marginal tax rates as data points for estimation of a marginal labor income tax function (the marginal tax rate as a function of labor income). We pool the annual marginal tax rates and estimate a single marginal labor income tax function. To facilitate the pooling, we deflate labor income to 2003 levels using the internal deflator from spell data wages described above. Furthermore, our structural model of worker search requires a marginal taxes as a function of hourly wages, not annual labor income. We transform the observed labor income into hourly wages by dividing by a measure of annual full-time working hours, as constructed by Lund and Vejlin (2015).[22]

As individual tax liabilities $T_t(LI_{it}, \mathbf{Z}_{it})$ are not solely functions of labor income LI, but also depend on other income components, e.g. capital income, and household characteristics, the mapping from hourly wages to marginal tax rates is not one-to-one. To proceed with estimation of the marginal labor income tax functions, we split hourly wages into 100 bins representing the percentiles in the empirical distribution of hourly wages for a given year in a given tax regime. We next compute the median marginal tax rates for each hourly wage bin in each year, pool the annual median marginal tax rates and fit a flexible parametric function (of hourly wages) to these median median marginal tax rates. Figure 9 superimpose the estimated marginal tax functions on the annual within-percentile median tax rates.

[22] According to this measure, annual hours range between 1,665 and 1,695 for the estimation period 1999-2003.

Figure 9: Marginal tax functions for Denmark (baseline years)



7 Estimation

We estimate the model by indirect inference. For now we focus on the just-identified case.

7.1 Auxiliary Model

We specify one moment for each of the structural parameters, which are

$$[s_i, A, c_1, y_l, y_h, \alpha_i, \beta_i, \tau_i, \sigma_\epsilon^2] \tag{23}$$

where $s_j, \alpha_j, \beta_j, \pi_j$ all are vectors of size J, where J is the number of types that we allow for. σ_ϵ is the variance of measurement error. We normalize $c_0 = 1$ as this cannot separately identified from A. We also set $y_0 = 0$ and $\gamma = 1$.

7.2 Notation

Let us define the variables used from the data

- $i = 1, ..., N$ index individuals
- $\ell = 1, ..., L_i$ index employment spells-a spell of consistent employment with no non-employment in between
- LC_i be an indicator variable for whether employment spell i is left-censored.

- $h, ..., H_{i\ell}$ index a job spell that occurs within employment spell ℓ for individual i
- $t = 1, ..., T_{i\ell h}$ index the set of wage observations on job spell $i\ell h$.
- $f_{i\ell h}$ the establishment associated with this job spell
- $D_{i\ell h}$ the duration of time that the worker worked on job spell $i\ell h$
- $w_{i\ell ht}$ the tth wage observation at job $i\ell h$
- E_{it} is the labor market status of individual i in the t'th period where the worker is present in the data. It takes the value 1 if the worker is employed and 0 otherwise.
- $t = 1, ..., T_i$ index the set of years, where the worker is present in the data.
- $k = 1, ...K_i$ the number of non-employment spells for individual i
- D^n_{ik} the duration of non-employment spell ik

We assume that the panel runs for T years.

Define the average wage for worker i as

$$\overline{w_i} \equiv \frac{\sum_{\ell=1}^{L_i} \sum_{h}^{H_{i\ell}} \sum_{t=1}^{T_{i\ell h}} w_{i\ell ht}}{\sum_{\ell=1}^{L_i} \sum_{h}^{H_{i\ell}} T_{i\ell h}}$$

and the average job wage in job $i\ell h$ as

$$\overline{w_{i\ell h}} \equiv \frac{\sum_{t=1}^{T_{i\ell h}} w_{i\ell ht}}{\sum_{\ell=1}^{L_i} T_{i\ell h}}$$

7.2.1 Auxiliary Parameters

In this section we define the auxiliary parameters that we try to match. For each structural parameters we chose one auxiliary and explain how it helps identify the structural parameter.

- The scale of the matching function, A: We propose to use the unemployment rate for those that are employed at some time during the panel. The reason we choose to condition on employment is that we otherwise would add a lot of workers that are outside the labor market.

$$\frac{\sum_{i=1}^{N}\sum_{t=1}^{T_i}(1-E_{it})}{\sum_{i=1}^{N} T_i 1[\sum_{t=1}^{T_i} E_{it} > 0]}$$

- The convexity of the search cost function, c_1: We suggest to use the variation in the duration of the first job. The basic idea is that the more convex the cost function is, the more search choices will be the same across workers, which will decrease the variance of the duration. We propose to use the first job out of unemployment, but we could have used any job number. We also restrict it to jobs that end in a job-to-job transition. This is because we want to avoid heterogeneity in s_j to affect the moment.

$$\frac{\sum_{i=1}^{N}\sum_{\ell=1}^{L_i-1}(1-LC_i)(D_{i\ell 1}-\overline{D_{i\ell 1}})^2}{\sum_{i=1}^{N}\sum_{\ell=1}^{L_i-1}(1-LC_i)}$$

where $\overline{D_{i\ell 1}} = \frac{\sum_{i=1}^{N}\sum_{\ell=1}^{L_i-1}(1-LC_i)D_{i\ell 1}}{\sum_{i=1}^{N}\sum_{\ell=1}^{L_i-1}(1-LC_i)}$.

- The limits on the wage distribution, y_l, y_h: We set these to match the 1 and then 99 percentile of the individual average wage distribution, $\overline{w_i}$.
- Measurement error in wages, σ_ϵ^2: Here we use within match variation in wages.

$$\frac{\sum_{\ell=1}^{L_i}\sum_{h}^{H_{i\ell}}\sum_{t=1}^{T_{i\ell h}}(w_{i\ell ht}-\overline{w_{i\ell h}})^2}{\sum_{\ell=1}^{L_i}\sum_{h}^{H_{i\ell}}T_{i\ell h}}$$

- Worker type specific job destruction rate, s_j: We use variation in the hazard rates out of employment across different percentiles of the worker average wage distribution. Specifically, we calculate

$$\frac{\sum_{i=1}^{N}\sum_{\ell=1}^{L_i} H_{i\ell-1}1[\overline{w_i}\in \mathbb{P}_{\daleth}]}{\sum_{i=1}^{N}\sum_{\ell=1}^{L_i}\sum_{h}^{H_{i\ell}} D_{i\ell h}1[\overline{w_i}\in \mathbb{P}_{\daleth}]}$$

where $\mathbb{P}_{\daleth}$ is the set of average worker wages in rung j, where the rungs are the intervals $[0, 1/J, ..., 1]$ of the CDF over average worker wages, w_i.

- The worker specific offer distributions, α_i, β_i: Here we use within worker-between match variation in wages conditional on the average worker wage. That is we target the 75th and 25th percentile in the distribution of $\overline{w_{i\ell h}} - \overline{w_i}$ conditional on the average wage of the worker, w_i, being in the interval of $[0, 1/J, ..., 1]$ of the

CDF over average worker wages, w_i conditional on having more than one job in the sample.

- The worker type distribution, τ_i: We suggest to use the between worker variation to pin down the type proportions. Basically, if we have a sufficiently long panel then the average wage will perfectly pin down the worker type. Since we need $J-1$ moments to identify the proportions, we suggest to use the differences between the $J/(J+1)$'th percentile and the $(J-1)/(J+1)$'th, the $(J-1)/(J+1)$'th and the $(J-2)/(J+1)$'th, and so forth.

7.3 Model Fit and Estimated Parameters

In this section we present some very preliminary structural parameter estimates relating to the model specification without amenities. The estimated structural model allows us to compare the actual tax schedule to the optimal schedule obtained from the estimated model. For this estimation exercise we consider a specification with only 2 worker-types and 10 firm-types, which implies that the model contains 10 free parameters. We estimate the these parameters by Indirect Inference.[23] We focus attention for now on the just-identified case, i.e. we are searching for 10 structural parameters to fit the 10 auxiliary statistics described above.

7.3.1 Structural parameter estimates

Table 3 contains the estimated values for the 12 structural parameters. Standard errors are not reported.

[23] In fact, the model contains 12 free parameters, but the support of the gross wage distribution is directly observable in the data, and we fix the lower and upper point of support in this way.

Table 3: Structural parameter estimates

Parameter	Description	Estimate/Value
r	Effective discount rate (fixed, not estimated)	0.05
A	Scale parameter in search technology	1.45
c_0	Scale, search cost function (normalized, not estimated)	1
c_1	Elasticity of search cost function	0.64
y_l	Lower bound, gross wage dist. (fixed prior to est.)	57
y_h	Upper bound, gross wage dist. (fixed prior to est.)	543
σ_ϵ	Standard deviation, log wage measurement error	0.11
s_1	Job destruction rate, worker type 1	0.11
s_2	Job destruction rate, worker type 2	0.02
α_1	Worker type 1, Offer Distr., Alpha Parameter	0.25
α_2	Worker type 2, Offer Distr., Alpha Parameter	2.01
β_1	Worker type 1, Offer Distr., Beta Parameter	1.53
β_2	Worker type 2, Offer Distr., Beta Parameter	3.98
τ	Share of type-1 workers	0.37
$\kappa_j = \kappa$	Worker type efficiency (fixed, not estimated)	1
$\bar{M}$	Amount to raise	256.52
γ	Constant relative risk aversion	1

Table 4 tabulates data moments and the model implied counterparts, computed at the estimated structural parameter values. Details on the computation of the auxiliary statistics are available in section 7.2.1.

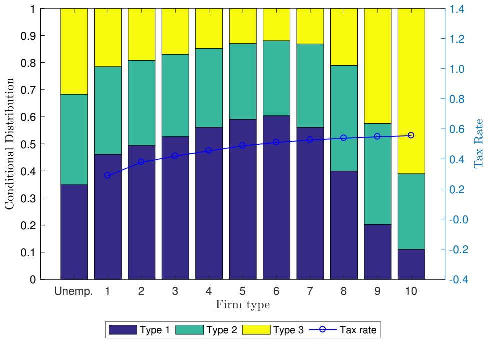
7.3.2 Analysis

Based on the structural parameters reported above representing the Danish 1999-2003 economy, we can re-do our analysis of the welfare effects of taxation, and compare the actually implemented tax system to that stipulated by our social planner, who trade off efficiency and equity in an optimal manner. In this part of the analysis, we compare four scenarios: The actual tax sytem, a Poll tax regime, and proportional tax regime and the optimal tax regime. The four cases are made comparable by ensuring that they raise the same amount of government revenue $\overline{M}$.[24]

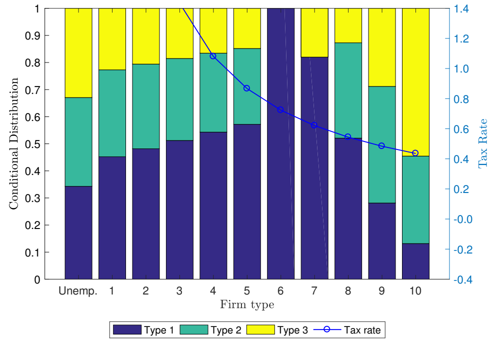
[24]In section 5 we conducted a similar illustrative analysis of four cases, No tax, Poll tax, proportional tax, and optimal tax. There, we imposed the government raise the same revenue in lifetime terms. Here, we

Table 4: Model fit

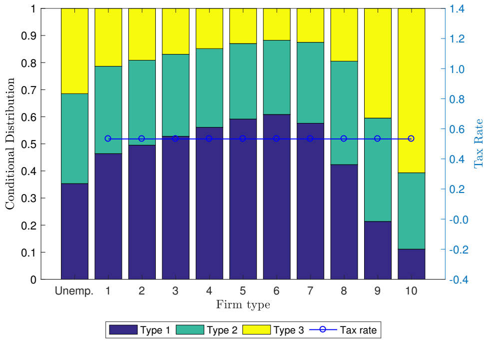
Auxiliary statistic	Data value	Model value
Cross section unemployment rate	0.1179	0.0604
Variance of the duration of first job out of unemployment	1.4559	1.8298
1st percentile, individual average wage distribution	57	57
99th percentile, individual average wage distribution	543	543
Within-match variance of log wages	0.0106	0.0106
Annual employment hazard rate $\mid$ wage $\leq$ median wage	0.2075	0.1633
Annual employment hazard rate $\mid$ wage $>$ median wage	0.0970	0.1157
25th percentile in within-worker, between-match dist. of $\ln w \mid$ worker avg. $\ln w \leq$ median	-0.0521	-0.4021
75th percentile in within-worker, between-match dist. of $\ln w \mid$ worker avg. $\ln w \leq$ median	-0.0467	0.0224
25th percentile in within-worker, between-match dist. of $\ln w \mid$ worker avg. $\ln w >$ median	0.0664	0.0415
75th percentile in within-worker, between-match dist. of $\ln w \mid$ worker avg. $\ln w >$ median	0.0538	0.3321
Difference between 67th and 33th percentile in worker avg. log wage dist.	0.2617	0.2123



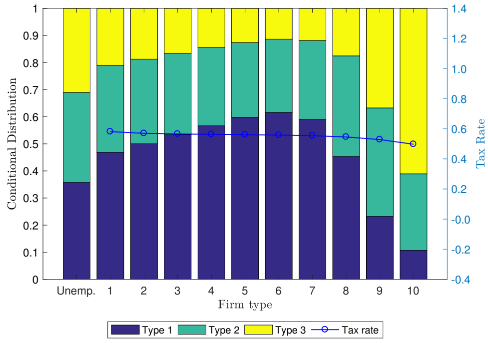
(a) Actual Tax System



(b) Poll Tax



(c) Proportional Tax



(d) Optimal Tax

Figure 10: State Conditional Worker Distribution

Figure 10d shows the resulting tax functions for the four cases. We note that, for these very preliminary estimates, the optimal tax system is actually slightly regressive, but overall resembles a proportional tax.

8 Conclusion

We have provided a theoretical and empirical study of the optimal taxation of labor income in the presence of search frictions and unobservable amenities using comprehensive Danish matched employer-employee data. With search frictions, on-the-job search, and heterogenous workers, income taxation impact the returns to search, and therefore have allocative consequences. The optimal tax system trade off efficiency against equity concerns at the margin. Equity concerns make it desirable to levy low taxes on (or indeed, subsidize) low-wage jobs including unemployment, and levy high taxes on high-wage jobs. Efficiency concerns limit how much taxes an optimal tax

fix the revenue in steady-state flow terms as computed by the estimated structural model under the actual Danish tax system.

system levy on high-paid jobs, as high taxes distort the workers' incentives to search. Using detailed micro data on wages, labor market transitions, and income tax filings we estimate the structural model and compare the the actual Danish tax regime with the optimal one. Preliminary results suggest the optimal tax schedule exhibits less progressivity than the actual tax system in place. The model allows us to quantify the welfare gains from adopting an optimal income tax schedule.

Several extensions on both the theoretical framework and the empirical analysis presented in this paper are underway. First of all, we want to incorporate amenities in our simulated examples, and eventuall, in the empirical analysis. Second, in a separate note (Bagger, Moen, and Vejlin, 2017) we extend the model and allow for directed search and entry of firms as in Moen (1997) and Garibaldi and Moen (2016) in order to incorporate the effect of taxes on wage formation. Third, we want to extend the model by assuming that the search costs are partly deductible for tax purposes. In this case, the tax system both reduces the costs and the gains from search, and the incentives to search will depend on the slope of the marginal tax rates.

References

Arseneau, D. M., and S. K. Chugh (2012): "Tax Smoothing in Frictional Labor Markets," *Journal of Political Economy*, 120(5), 926–985.

Bagger, J., M. Hejlesen, K. Sumiya, and R. M. Vejlin (2017): "Equilibrium Labor Allocation and Income Taxation," mimeo.

Bagger, J., and R. Lentz (2018): "An Empirical Model of Wage Dispersion with Sorting," *Review of Economic Studies*.

Bagger, J., E. R. Moen, and R. M. Vejlin (2017): "Competitive On-the-Job Search with Taxes and Amenities," mimeo.

Best, M., and H. Kleven (2013): "Optimal income taxation with career effects of work effort," mimeo.

Burdett, K., and D. T. Mortensen (1998): "Wage Differentials, Employer Size, and Unemployment," *International Economic Review*, 39(2), 257–273.

Chetty, R. (2009): "Sufficient Statistics for Welfare Analysis: A Bridge Between Structural and Reduced-Form Methods," *Annual Review of Economics*, 1, 451–488.

Christensen, B. J., R. Lentz, D. T. Mortensen, G. R. Neumann, and A. Werwatz (2005): "On-the-Job Search and the Wage Distribution," *Journal of Labor Economics*, 23(1), 31–58.

Christiansen, V. (2015): "Optimal Participation Taxes," *Economica*, 82(595-612).

Diamond, P. A. (1982): "Wage Determination and Efficiency in Search Equilibrium," *Review of Economic Studies*, 49(2), 217–227.

Domeji, D. (2005): "Optimal Capital Taxation and Labor Market Search," *Review of Economic Dynamics*, 8(3), 623–650.

Garibaldi, P., and E. R. Moen (2016): "Competitive On-the-Job Search," *Review of Economic Dynamics*, 19, 88–107.

Geromichalos, A. (2015): "Unemployment Insurance and Optimal Taxation in a Search Model of the Labor Market," *Review of Economic Dynamics*, 18, 365–380.

Golosov, M., P. Maziero, and G. Menzio (2013): "Taxation and Redistribution of Residual Income Inequality," *Journal of Political Economy*, 121(6), 1160–1240.

Heathcote, J., K. Storesletten, and G. L. Violante (2017): "Optimal Tax Progressivity: An Analytical Framework," *Quarterly Journal of Economics*, 132(4), 1693–1754.

Hungerbuhler, M., E. Lehmann, A. Parmentier, and B. van der Linden (2006): "Optimal Redistributive Taxation in a Search Equilibrium Model," *Review of Economic Studies*, 73(3), 743–767.

Jiang, W. (2012): "Tax Policy in a Search and Matching Model with Heterogeneous Agents," mimeo.

Kleven, H. J., and E. A. Schultz (2014): "Estimating Taxable Income Responses Using Danish Tax Reforms," *American Economic Journal: Economic Policy*, 6(4), 271–301.

Lund, C. G., and R. Vejlin (2015): "Documenting and Improving the Hourly Wage Measure in the Danish IDA Database," *Danish Journal of Economics.*

Moen, E. R. (1997): "Competitive Search Equilibrium," *Journal of Political Economy*, 105(2).

Mortensen, D. T. (1982): *The Economics of Information and Uncertainty,* chap. 7, pp. 233–258. University of Chicago Press.

Pissarides, C. (1985): "Short-run Equilibrium Dynamics of Unemployment, Vacancies, and Real Wages," *American Economic Review*, 75(4), 676–690.

——— (1994): "Search Unemployment with On-the-Job Search," *Review of Economics Studies*, 61(3), 457–475.

Saez, E. (2002): "Optimal Income Transfer Programs: Intensive versus Extensive Labor Supply Responses," *Quarterly Journal of Economics*, 117(3), 1039–1073.

Shi, S., and Q. Wen (1999): "Labor Marekt Seach and the Dynamic Effects of Taxes and Subsidies," *Journal of Monetary Economics*, 43(2), 457–495.

Sleet, C., and H. Yazici (2017): "Taxation, Redistribution and Frictional Labor Supply," mimeo.

Wilemme, G. (2017): "Optimal Taxation to Correct Job Mismatching," mimeo.