I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 14865

Equilibrium Worker-Firm Allocations and the Deadweight Losses of Taxation

Jesper Bagger
Espen R. Moen
Rune Vejlin

NOVEMBER 2021

I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 14865

Equilibrium Worker-Firm Allocations and the Deadweight Losses of Taxation

Jesper Bagger
University of London and Aarhus University

Espen R. Moen
BI Norwegian Business School

Rune Vejlin
Aarhus University and IZA

NOVEMBER 2021

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Equilibrium Worker-Firm Allocations and the Deadweight Losses of Taxation

We analyse the deadweight losses of tax-induced labor misallocation in an equilibrium model of the labour market where workers search to climb a job ladder and firms post vacancies. Workers differ in abilities. Jobs differ in productivities and amenities. A planner uses affine tax functions to finance lump-sum transfers to all workers and unemployment benefits. The competitive search equilibrium maximizes after-tax utility subject to resource constraints and the tax policy. A higher tax rate distorts search effort, job ranking and vacancy creation. Distortions vary on the job ladder, but always result in deadweight losses. We calibrate the model using matched employer-employee data from Denmark. The marginal deadweight loss is 33 percent of the tax base, and primarily arise from distorted search effort and vacancy creation. Steeply rising deadweight losses from distorted vacancy creation imply that the deadweight loss in the calibrated economy exceeds those incurred by very inequality averse social planners.

JEL Classification: H21, H30 J63, J64

Keywords: deadweight loss, optimal taxation, redistribution, labour allocation, job search, job ranking, vacancy creation, amenities, matched employer-employee data

Corresponding author:
Rune Vejlin
Department of Economics and Business Economics
Aarhus University
Fuglesangs Alle 4
8210 Aarhus V
Denmark

E-mail: rvejlin@econ.au.dk

1 Introduction

The equilibrium allocation of resources does not materialize costlessly in markets with frictions. Bringing together buyers and sellers in such markets is value creation that may be distorted by taxation. In a frictional labour market, income taxation impacts the unemployment rate and the allocation of workers to firms, and gives rise to hitherto understudied deadweight losses.

We study these deadweight losses of taxation in a rich model of a frictional labour market. Unemployed workers search to find a job, employed workers search on-the-job to locate a better job, and firms are free to enter the market and search for employees by posting vacancies. Workers differ in intrinsic ability (a worker's type). Jobs differ in their productivities and in the level of amenities they provide. These amenities are not observed by the government and we assume that they cannot be taxed. Workers rank jobs on a job ladder with unemployment at the bottom rung, followed by unattractive jobs at lower rungs and more attractive jobs at higher rungs. Workers exert search effort to climb the job ladder. We consider affine labour income tax functions, i.e., a proportional tax on labour income that finances lump-sum transfers to all workers, unemployment benefits, and exogenous expenditures.

In competitive search equilibrium, workers of different types and at different rungs of the job ladder search in separated submarkets and do not create search externalities for each other. We assume employment contracts are sufficiently sophisticated to resolve any employer-employee agency problems. The tax and benefit system, however, introduces two fiscal externalities: workers and firms internalize neither their impact on the tax base, and thus on the government's tax revenue, nor on the government's unemployment benefit expenditures. In a laissez-faire economy without taxes and benefits, the competitive search equilibrium coincides with the planner's solution. With taxes and benefits, the equilibrium maximizes the expected after-tax lifetime utility of searching workers, but distorts the allocation of workers to jobs away from the planner solution. Indeed, an increase in the tax rate impacts workers' job search effort, their ranking of jobs, and firms' vacancy creation. These distortions generate deadweight losses.

Consider first distortions to search effort. As workers' search costs are not deductible, a tax increase reduces workers' search effort on job ladder rungs where taxable incomes are expected to increase after successful search. This is the case for unemployed workers, and a higher income tax rate therefore leads to lower search effort among the unemployed and a deadweight loss: the tax base shrinks and aggregate unemployment benefit expenditures increase. For workers employed in low-productivity jobs, whose wages also tend to increase as they climb the job ladder, search effort fall as well, further shrinking the tax base. In effect, at these job ladder rungs, income taxation gives rise to a hold-up problem: the worker carries the entire cost of search, while the government expropriates part of the return. Workers in high-productivity, low-amenity jobs, however, expect wage incomes to fall as they move up the job ladder, as, empirically, wages do for a substantial minority of job switchers. At these job ladder rungs, income taxation effectively subsidizes search effort, and higher taxes therefore spur search effort. This nonetheless generates a deadweight loss because it increases the rate at which workers leave high-paying, high-productive jobs for lower-paying, high-amenity jobs.

Next, consider distortions to workers' ranking of jobs. A worker accepts employment at a new employer if and only if the new job offers better terms of employment than the worker's current job (or unemployment, as the case may be). Since wages are taxed while amenities are not, income taxation tend to make high-amenity, low-productivity jobs more attractive relative to high-productivity, low-amenity jobs. Relative to the stipulations of a social planner, income taxation in the decentralized equilibrium induces workers to accept too few high-productivity, low-amenity jobs and too many low-productivity, high-amenity jobs, and an increase in the tax rate strengthen these tendencies, which reduces the tax base and gives rise to deadweight losses.

Finally, consider distortions to vacancy creation. More vacancies imply that workers and firms match faster, but also higher vacancy costs for firms. A zero-profit condition implies that these costs are ultimately borne by the workers through lower wages after successful search, creating a trade-off between the job finding rate and the expected wage after finding a job. In competitive search equilibrium, this trade-off is balanced to maximize the expected after-tax NPV utility of searching workers. In the absence of amenities, higher expected income due to a higher job finding rate and higher expected income due to higher wages when finding a job are taxed at the same rate. In that case, the equilibrium trade-off between high job finding rates and high wages is unaffected by taxation, and vacancy creation is left undistorted.

However, when the economy features untaxed amenities, a tax increase reduces the gains from higher wages proportionally, but reduces the gains from a higher job finding rates less or more than proportionally depending on whether the amenity level is expected to increase or decrease along the job ladder. In submarkets where workers expect amenity increases, e.g. in high-productivity low-amenity submarkets, a higher job finding rate offers faster access to these tax-free amenities. Hence, in such submarkets, a tax increase leads to excessive vacancy creation and worker reallocation, which shrinks the tax base and gives rise to deadweight losses, except at the unemployment rung; here, the excessive vacancy creation in fact counteracts the depressing effect of unemployment benefit provision on vacancy creation, and therefore reduces the deadweight loss. At job ladder rungs where workers expect negative amenity growth, e.g. in low-productivity high-amenity submarkets, deadweight losses arise because a tax increase depresses vacancy creation and labour reallocation.

Our paper also offers an important methodological contribution. We construct an equilibrium on-the-job search model of worker reallocation that is sufficiently rich to admit detailed qualitative and quantitative analysis of interesting policy questions, yet delivers an allocation that maximizes after-tax utility, and is constrained efficient in laissez-faire. The latter feature has several advantages. First, it offers clarity of interpretation in the sense that deadweight losses can be traced directly to the fiscal externalities of the tax and transfer system; specifically, tax distortions are unaffected by externalities arising from arbitrary assumptions regarding the matching technology and the employment contracts. Second, it offers analytical tractability; specifically, partial deadweight losses can be derived independently at each job ladder rung using the Envelope Theorem, while the model's recursive structure implies that the total deadweight loss obtains by integrating the rung-specific (partial) losses. Third, it implies that the (marginal) deadweight losses come about via deflated tax bases and inflated unemployment benefit expenditures, with the former component being proportional to the tax rate. This connects our paper to the broader literature on the deadweight loss of taxation and on optimal taxation.

For the quantitative part of our analysis, we calibrate the model parameters using matched employer-employee data from Denmark for 1994-2003, including detailed information on individual tax filings. We use the calibrated model to compute the implied marginal deadweight loss from (a linearised version of) the Danish tax and transfer system, and decompose the marginal deadweight loss into the three channels discussed above: job search effort, job ranking, and vacancy creation. The average marginal tax rate in Denmark is 0.643. We find that the marginal deadweight loss of income taxation is non-trivial and is 33 percent of the tax base. The elasticity of the tax base with respect to the income tax rate, an important component of our optimal tax formula, is 0.234. The deadweight loss arises because the tax and transfer system skews the allocation of workers away from high-productivity, low-amenity jobs and towards low-productivity, high-amenity jobs. Distortions to job search effort, job ranking, and vacancy creation comprises 28 percent, 7 percent, and 65 percent of the marginal deadweight loss, respectively.

Finally, we compute the optimal tax rates as a function of a social planner's aversion to inequality. The deadweight loss in the calibrated economy is substantially higher than the deadweight loss even very inequality averse planners are willing to incur to achieve their redistributive goals. We show that steeply rising deadweight losses coming from distorted vacancy creation are behind this finding. Careful account of exogenous government spending in the calibration may help rationalize the observed tax rate, but it remains the case that government spending entails a high deadweight loss, which primarily arises from distorted vacancy creation.

Related literature. A few papers in the early literature on search in the labour market consider the effects of taxation on the equilibrium allocation. Pissarides (1985) and Smith (1994), and somewhat later Pissarides (1998), analyse the effects of taxes on the unemployment rate from a positive perspective (see Pissarides (2000) for an overview).

In a bargaining framework, search frictions may create externalities, and a series of papers analyse how taxation may internalize these externalities. Boone and Bovenberg (2002) show how taxes may restore efficiency in equilibrium models of random search when the Hosios condition is violated.[1] Arseneau and Chugh (2012) studies taxation in a calibrated DSGE model with search frictions, and argues that cyclical variations in the search-based labour wedge call for taxes that vary over the business cycle. Wilemme (2021) studies taxation in a model of mismatch, and shows that taxes should be regressive to correct for workers not being sufficiently selective.

There is also a literature that studies optimal redistributive taxation and search when the planner has equity concerns. This literature is mostly concerned with the search decisions of

[1]Another set of papers analyse the role of taxes for wealth accumulation among workers in models where unemployed workers search to find a job. Shi and Wen (1999) analyse the effect of taxes in a model of random unemployed search, in which workers accumulate capital. Higher labour taxes discourage working, and lead to lower investments by firms and lower wages. Capital taxation on the other hand increases labour supply, as workers get a lower return on their capital. Hence, capital taxation may improve the allocation of resources. Domeji (2005) analyses optimal taxation within the same modelling framework, and find that the optimal capital tax is zero if and only if the Hosios's condition is satisfied. Jiang (2012) uses a similar setup to analyse the welfare effects of a UK tax reform.

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 we do) that workers are risk neutral, while the planner has preferences over the (expected) income distribution over different worker types. A revelation mechanism can be applied at the bargaining stage, so that the worker and the firm bargain over what "worker type" to reveal to the planner. As 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 who 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. Workers face a fixed cost 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, unemployment insurance makes workers indifferent between searching for any job and not searching, as this result in maximum insurance given workers' incentive compatibility constraint. There are no transfers between workers searching for different firm types; firms in effect finance the unemployment benefits of the workers they attract but do not hire. As a result, optimal taxation is regressive. Geromichalos (2015) studies optimal taxation with risk averse workers in a one-shot urn ball model and finds that unemployment benefits financed by lump-sum taxes lead to inefficiently high wages and low firm entry.

Kreiner, Munch, and Whitta-Jacobsen (2015) study the effects of taxes on workers' on-the-job search effort, and our work has some overlap with theirs. They work, however, with a one-sided search model, with a fixed arrival rate of jobs, wages equal to productivity, and no wage posting by firms. That is, their model does not feature equilibrium feedback from the firm side, as in a Diamond-Mortensen-Pissarides search model like ours, and they do not consider amenities.[2]

A couple of recent papers analyse taxation and on-the-job search in two-sided search models. Sleet and Yazici (2017) study optimal taxation in a Burdett and Mortensen (1998) model of on-the-job search. A tax on labour income (but not benefits) reduces net wages for workers, increases their before-tax reservation wage, and shifts the entire wage distribution and workers' production effort. This matters for the design of optimal tax policy because workers are risk averse and unable to smooth consumption. Our analysis is complementary. First, we analyse deadweight losses along the search effort, job ranking, and vacancy creation margins, which are not present in Sleet and Yazici's paper. Indeed, on-the-job search has no allocative consequences in their model, it only influences the division of rent. In contrast, we focus on distortions that influence the speed and direction of worker flows, and thus the worker-firm allocation. Second, optimal tax policy in Sleet and Yazici (2017) depends crucially on wages being set by bargaining where the firm has all the bargaining power. This is arguably arbitrary and fails to deliver constrained efficiency in the absence of taxation. Hence, in their model, taxation may have beneficial effects by mitigating inefficiencies introduced by the assumed wage setting mechanism. Our analysis

[2] Mancino and Mullins (2020) study the Earned Income Tax Credit (in the US) also using a one-sided on-the-job search model, but with hours constraints and multiple job holding.

is based on a constrained efficient model where deadweight losses can be attributed to market frictions as such. We think this is a considerable advantage, but it makes it challenging to go beyond affine tax functions and to study general tax functions.

Bagger, Hejlesen, Sumiya, and Vejlin (2017) estimate an equilibrium on-the-job search model with Burdett and Mortensen (1998) wage setting and two-sided endogenous search effort to measure the (long run) elasticity of taxable labour income (ETLI), and to evaluate Danish income tax reforms. The analysis in Bagger, Hejlesen, Sumiya, and Vejlin (2017) is entirely positive: it aims to accurately quantify the effect of non-linear income taxation and actual income tax reforms in an economy with search frictions and inefficiencies unrelated to taxes and transfers arising from the assumed matching technology and equilibrium wage setting game. We maintain an interest in quantitative predictions, but our focus is normative: we characterize deadweight losses (and the optimal redistributive tax) in a constrained efficient economy with search frictions.[3] Moreover, Bagger, Hejlesen, Sumiya, and Vejlin (2017) do not consider the role of amenities in guiding worker reallocation

Our paper also relate to papers outside the search literature. For example, 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: In Saez's model, reducing taxes at a given level induces some workers who were previously choosing an occupation one level above or below to switch to that level. In our model, by contrast, reducing taxes at a given job ladder rung reduces the search incentives for workers at that rung, increases search incentives for all lower-rung workers, and leaves search incentives unchanged for all higher-rung workers.

Finally, our work contributes to understanding how amenities shape outcomes in labor markets with frictions. The literature on search and amenities goes back to Hwang, Mortensen, and Reed (1998). More recently, Sorkin (2018) finds that amenities comprise over 50 percent of the firm component in the variance of earnings. Hall and Mueller (2018) finds that the standard deviation of offered wages is smaller (0.24 US Dollars) than that of the offered non-wage component (0.34 US Dollars), and Taber and Vejlin (2020) finds that the variance of flow utility is around 2.5 times higher than the variance of wages.

2 The Job Ladder Model

There is a unit mass of infintely lived risk neutral workers who discount the future at rate r, and who can be one of I worker types. The fraction of type-i workers is κ_i such that $\sum_{i=1}^{I} \kappa_i = 1$.

A worker is either unemployed or employed. Both unemployed and employed workers search

[3] Breda, Haywood, and Wang (2019) studies the effects of payroll taxation and minimum wage policies (in France) using an equilibrium on-the-job search model with sequential auction wage setting. The equilibrium is inefficient, and their focus is (largely) on the positive analysis of payroll tax reductions.

for jobs. Search effort is denoted e and is chosen by workers subject to a utility cost, $c(e)$, with $c'(e) \geq 0$ and $c''(e) > 0$. That is, job search is associated with disutility, not reduced income, and the cost of search is independent of the worker's type and income. The latter assumption implies that search effort depends on utility differentials in the available jobs, but not the level of utility in these jobs. This is a standard assumption in the labour search literature and is also in line with e.g. Saez (2002) who assumes that the choice of sectors do not depend on the income levels in the different sectors, only the differences in income between them.

Identical firms enter the economy at cost K, and are subject to profit taxation at flat rate τ. A fraction γ_K of the entry costs is deductible from profit taxation, so the net entry cost is $K(1-\gamma_K\tau)$. After entry, the firm is in possession of one vacancy. The flow cost of operating the vacancy is c^v with a fraction γ_c being deductible. When the firm finds a worker, the vacancy is immediately re-posted. Thus, firms have multiple jobs and there is no opportunity cost of hiring. Firms never exit the market and discount the future at the same rate, r, as workers.[4]

Upon meeting, a firm and a type-i worker draws a two-dimensional vector of job attributes $\mathbf{y} = (y_p, y_z) \in \mathscr{Y} \subset \mathbb{R}^2$ in the worker-type specific sampling distribution F^i. The first job attribute, y_p, is the productivity of the match. The second attribute, y_z, is the amenity of the job, observable to both parties and consumed by the worker. We refer to a job with attributes $\mathbf{y}$ as a type-$\mathbf{y}$ job. After the job attributes are realized, which also specifies the before-tax flow wage w to be paid to the worker, the parties decide whether or not to form a match. We assume that the productivity and amenity attributes are continuously distributed, that the sampled amenity attributes are independent of worker-type, and that the productivity and amenity attributes are sampled independently. That is, if f^i is the joint sampling density of $\mathbf{y}$ for type-i workers, then $f^i(\mathbf{y}) = f_p^i(y_p)f_z(y_z)$, where f_p^i and f_z are the marginal sampling densities of y_p and y_z, respectively. Higher type-i workers face better sampling distributions of productivity attributes (in a stochastic dominance sense).

The planner observes profits and wages, which is equivalent to observing the productivity attribute, y_p, but does not observe y_z. Hence, taxes can not be made contingent on y_z; indeed, the planner levies income taxes using the affine tax function $t(w) = tw - t_0 r$, where t is the proportional tax rate and t_0 is the net present value (NPV) lump-sum tax/transfer. Hence, the after-tax utility flow to an employed worker earning a before-tax wage w is $w - t(w) + y_z$, while the employing firm's after-tax profit flow is $(y_p - w)(1-\tau)$.

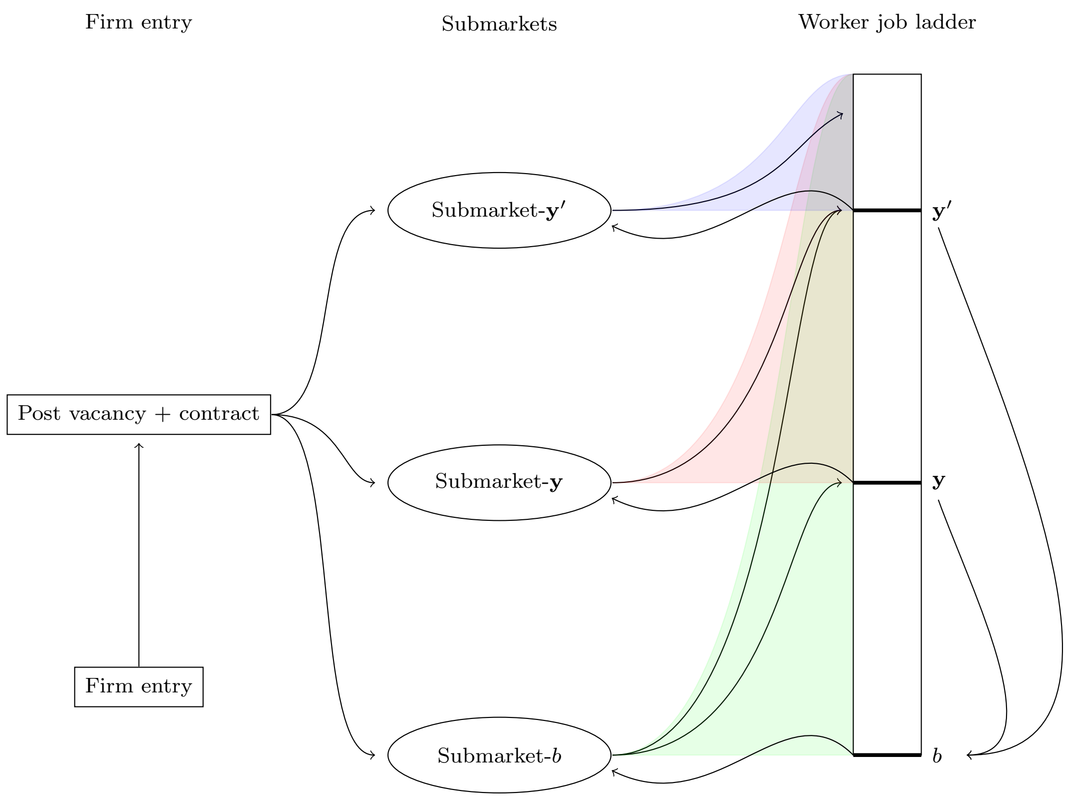
Jobs are destroyed at exogenous rate s at which point the worker initiates an unemployment spell. During unemployment, the worker receives a flow income transfer b irrespective of type, and enjoys amentity attribute $y_{z,0}$, which includes non-taxed home production. For convenience, we let $\mathbf{y}_0 = (b, y_{z,0})$ indicate the unemployment attributes. Unemployment benefits are subject to income taxation, so the utility flow of an unemployed worker is $b - t(b) + y_{z,0}$.

In competitive search equilibrium, firms post employment contracts. A contract specifies the remuneration of the worker that is hired, and this worker's search behaviour and job acceptance

[4]The assumption that workers and firms live forever simplifies the analysis slightly. Alternatively we may assume that r consists of a pure discount rate r and an attrition rate of workers, s_w, and that firms exit the market at rate, λ.

decision as an employee in the new job. The labour market endogenously separates into submarkets. In each of these submarkets, firms offer identical contracts, and workers are of the same type and work in jobs with the same job attributes (or are all unemployed); that is, submarkets are indexed by $(i, \mathbf{y})$. In each submarket, the flow of new matches is given by a Cobb-Douglas matching function $m(E, V) = AE^{\beta}V^{1-\beta}$, where E and and V are aggregate search effort by workers and the number of vacancies posted in the submarket. Let $\theta = \frac{V}{E}$ denote labour market tightness in a submarket, $p(\theta)$ the arrival rate of job offers to workers per unit of search effort, and $q(\theta)$ the arrival rate of workers to a vacancy. The elasticity of the job offer arrival and the vacancy filling rate with respect to tightness are $1-\beta$ and $-\beta$, respectively. The model structure is illustrated in Figure 1.

Figure 1: Model overview



Notes: The shaded green, red, and blue areas illustrate that workers reallocate to higher rungs when the opportunity arises. Rung-$\mathbf{y}$ contracts stipulate identical search strategies for all rung-$\mathbf{y}$ workers, but these workers are paid differently because they came to rung-$\mathbf{y}$ from different (lower-ranked) rungs.

2.1 Asset Value Equations

Consider an employed type-i worker currently matched with a type-$\mathbf{y}$ job. The key endogenous objects, search effort e, the set of job attributes that the worker accepts upon meeting a new vacancy, a set we denote $\mathscr{Y}_a \subseteq \mathscr{Y}$, the labour market tightness θ faced by the worker, and the wage w will be functions of the worker-type i and the attributes of the current job, $\mathbf{y}$, and, in the case of w, the attributes of the previous job (or unemployment, as the case may be), which we denote $\mathbf{y}^{\ell}$. However, for notational simplicity, we state asset values as functions only of i, $\mathbf{y}$, and when relevant, $\mathbf{y}^\ell$, and suppress the dependence of e, $\mathscr{Y}_a$, θ and w on i, $\mathbf{y}$ and $\mathbf{y}^\ell$.

Let $V^i(\mathbf{y},\mathbf{y}^\ell)$ denote the NPV utility of a type-i worker in a type-$\mathbf{y}$ job, who previously worked in a type-$\mathbf{y}^\ell$ job ($\mathbf{y}^\ell = \mathbf{y}_0$ if the worker was hired into the type-$\mathbf{y}$ job from unemployment). Furthermore, let V_0^i be the NPV utility of unemployment to a type-i worker, which is history independent. Since $V^i(\mathbf{y},\mathbf{y}^\ell)$ and V_0^i are an utilities, they describe the worker's after-tax situation. Omitting the lump-sum transfer t_0, the Bellman equation for $V^i(\mathbf{y},\mathbf{y}^\ell)$ is:[5]

$$(r+s)V^i(\mathbf{y},\mathbf{y}^\ell) = w - tw + y_z - c(e) + sV_0^i + ep(\theta)\mathrm{E}^{\mathbf{y}'\in\mathscr{Y}_a}\left[V^i(\mathbf{y}',\mathbf{y}) - V^i(\mathbf{y},\mathbf{y}^\ell)\right], \quad (1)$$

where $\mathrm{E}^{\mathbf{y}'\in\mathscr{Y}_a^i}\left[V^i(\mathbf{y}',\mathbf{y}) - V^i(\mathbf{y},\mathbf{y}^\ell)\right] \equiv \int_{\mathscr{Y}_a}\left[V^i(\mathbf{y}',\mathbf{y}) - V^i(\mathbf{y},\mathbf{y}^\ell)\right]dF(\mathbf{y}')$ is the expected gain from meeting a vacancy, which also introduces the notation $d\mathbf{y}' \equiv dy_p'\,dy_z'$. The permanent utility flow to an employed worker is the utility flow in the current job, plus the expected capital loss from job destruction, plus the expected capital gain from meeting a new vacancy.

Let $J^i(\mathbf{y},\mathbf{y}^\ell)$ denote the after-tax net present discounted income to the firm from a type-$\mathbf{y}$ job occupied by a type-i worker who was poached from a type-$\mathbf{y}^\ell$ job. Then

$$(r+s)J^i(\mathbf{y},\mathbf{y}^\ell) = (y_p - w)(1-\tau) - ep(\theta)\mathrm{E}^{\mathbf{y}'\in\mathscr{Y}_a}J^i(\mathbf{y},\mathbf{y}^\ell) \quad (2)$$

which states that the permanent profit flow to the firm from a filled job equals the current profit flow, plus the expected (negative) capital gain incurred when the worker quits the job.

The joint after-tax value of a matched worker and firm plays a key role in our analysis. The match value, which is measured in worker utils, is denoted $L^i(\mathbf{y})$, and is given by

$$L^i(\mathbf{y}) \equiv V^i(\mathbf{y},\mathbf{y}^\ell) + \frac{1-t}{1-\tau}J(\mathbf{y},\mathbf{y}^\ell), \quad (3)$$

for $\mathbf{y} \neq \mathbf{y}_0$. We further define $L_0^i \equiv V_0^i$ for unemployment. The worker-utility denoted after-tax match value reflects that firm after-tax profit can be exchanged for worker utility at rate $(1-t)/(1-\tau)$. Equations (1), (2), and (3) imply that

$$(r+s)L^i(\mathbf{y}) = y_p(1-t) + y_z - c(e) + sL_0^i + ep(\theta)\mathrm{E}^{\mathbf{y}'\in\mathscr{Y}_a}\left[V^i(\mathbf{y}',\mathbf{y}) - L^i(\mathbf{y})\right]. \quad (4)$$

Our notation indicates that $L^i(\mathbf{y})$ is independent of the worker's labour market history $\mathbf{y}^\ell$. Hence, $(r+s)L^i(\mathbf{y})$, comprises the flow utility, including the utility cost of worker search, as well as expected capital loses and gains associated with job destruction and meeting a vacancy.

Let $\pi^i(\mathbf{y})$ denote the expected income flow (including firms search cost) generated by a vacancy posted in the $(i,\mathbf{y})$-submarket. Then,

$$\pi^i(\mathbf{y}) = q(\theta)\mathrm{E}^{\mathbf{y}'\in\mathscr{Y}_a}J^i(\mathbf{y}',\mathbf{y}). \quad (5)$$

Since search is competitive, each submarket yields the same profit, so $\pi^i(\mathbf{y}) = \pi$ for all $(i,\mathbf{y})$. Thus, the NPV profit accruing to a firm of entering any markets is given by

[5]The omission of t_0 from (1) has no bearing on the behaviour of agents in the model, but does of course affect their welfare. We re-introduce t_0 when we conduct our welfare analysis.

$$\Pi = \frac{-(1-\gamma_c \tau)c^v + \pi}{r}.$$

There is a cost K to enter, a fraction γ_K of which is deductible, meaning that the net-of-tax entry cost is $K(1-\gamma_K\tau)$. Free entry ensures that firms enter up until the point where future expected net-of-tax profits exactly offsets the net-of-tax entry cost; that is, entry occurs up until $\Pi = K(1-\gamma_K\tau)$. This pins down the profit flow requirement, $\pi = \overline{\pi}$, where $\overline{\pi}$ is defined as

$$\overline{\pi} = rK(1-\gamma_K\tau) + (1-\gamma_v\tau)c^v. \tag{6}$$

For a given tax system, $\overline{\pi}$ is exogenous.

A few additional useful variables and relationships. Equations (1), (2), (5), and (6) suffices for establishing the competitive search equilibrium in our job ladder economy. However, to analyse the effects of income taxation, we need a few additional variables, relationships and representations that are naturally introduced here. In doing so, to further simplify our notation. We suppress the index i for worker type, and define the operator Δ as

$$\Delta X(\mathbf{y}) \equiv \mathrm{E}^{\mathbf{y}'\in\mathscr{Y}_a}\left[X(\mathbf{y}') - X(\mathbf{y})\right],$$

for any X. That is, $\Delta X(\mathbf{y})$ is shorthand for the expected gain in $X(\mathbf{y})$ from meeting a vacancy.

First, it proves convenient to substitute the zero-profit condition into the expression for $L(\mathbf{y})$ given by (4) to obtain an alternative representation of the joint match NPV utility:

$$(r+s)L(\mathbf{y}) = y_p(1-t) + y_z - c(e) + sL_0 + ep(\theta)\Delta L(\mathbf{y}) - \frac{1-t}{1-\tau}e\theta\overline{\pi}. \tag{7}$$

The before-tax NPV wage of a worker employed in a job with attributes $\mathbf{y}$, who entered the type-$\mathbf{y}$ submarket from the type-$\mathbf{y}^\ell$ submarket, is given by

$$(r+s)W(\mathbf{y},\mathbf{y}^\ell) = w + sW_0 + ep(\theta(\mathbf{y}))\Delta W(\mathbf{y},\mathbf{y}^\ell) \tag{8}$$

where $W_0 \equiv W(\mathbf{y}_0,\mathbf{y}_0)$ and w is replaced by b if $\mathbf{y} = \mathbf{y}_0$. A worker's NPV tax liability is $tW(\mathbf{y},\mathbf{y}^\ell)$ and in the analysis to come, we refer to $W(\mathbf{y},\mathbf{y}^\ell)$ and W_0 as the tax bases.

Let B_0 be the NPV unemployment benefits of an unemployed worker. It follows that

$$B_0 = \frac{(r+s)b}{r[r+s+ep(\theta)]}, \tag{9}$$

where e and θ in (9) refers to search effort and tightness in the unemployment submarket. Let B_1 be the NPV unemployment benefits of an employed worker. We have $B_1 = \frac{s}{r+s}B_0$.

We will also need to keep track of the joint NPV utility of a worker-firm pair before taxes and unemployment benefits. Define $M(\mathbf{y}) \equiv V(\mathbf{y},\mathbf{y}^\ell) + J(\mathbf{y},\mathbf{y}^\ell) + tW(\mathbf{y},\mathbf{y}^\ell) + T^\tau(\mathbf{y},\mathbf{y}^\ell) - B_1$, where $T^\tau(\mathbf{y}, \mathbf{y}^\ell)$ is NPV of profit taxes. It follows that

$$(r+s)M(\mathbf{y}) = y_p + y_z - c(e) + sM_0 + ep(\theta)\Delta M(\mathbf{y}) - e\theta(rK + c^v). \tag{10}$$

The last term in (10) reflects the real cost of operating a vacancy, which is either borne by the firm or the government depending on deductibility. Since benefits are not included in M, $M(\mathbf{y}_0)$ is defined slightly different, specifically with $y_p = 0$. With this in mind, define $M_0 = M(\mathbf{y}_0)$.

Finally, we want to separate L into taxable and non-taxable components. To that end, define

$$(r+s)Y_p(\mathbf{y}) = y_p + ep(\theta)\Delta Y_p(\mathbf{y}) - e\theta\frac{\overline{\pi}}{1-\tau} + sY_{p,0}, \tag{11}$$

$$(r+s)Y_z(\mathbf{y}) = y_z - c(e) + ep(\theta)\Delta Y_z(\mathbf{y}) + sY_{z,0}, \tag{12}$$

where $Y_{p,0} \equiv Y_p(\mathbf{y}_0)$ and $Y_{z,0} \equiv Y_z(\mathbf{y}_0)$. Inserting these into (7) reveals that

$$\begin{aligned}(r+s)L(\mathbf{y}) = (1-t)y_p + y_z - c(e) + s\left[(1-t)Y_{p,0} + Y_{z,0}\right] \\ + ep(\theta)\left[(1-t)\Delta Y_p(\mathbf{y}) + \Delta Y_z(\mathbf{y})\right] - \frac{1-t}{1-\tau}e\theta\overline{\pi}.\end{aligned} \tag{13}$$

Note that $L(\mathbf{y}) = (1-t)Y_p(\mathbf{y}) + Y_z(\mathbf{y})$. Hence, we can decompose the joint utility created by a match into a taxable component, $(1-t)Y_p(\mathbf{y})$, and a non-taxable component, $Y_z(\mathbf{y})$. The taxable component, $(1-t)Y_p(\mathbf{y})$, is the before-tax NPV of climbing the productivity-attribute ladder, net of the cost of creating the job ladder. The non-taxable component, $Y_z(\mathbf{y})$, is the NPV of climbing the amenity-ladder, net of the disutility from search.

3 Competitive Search Equilibrium

In competitive search equilibrium, the labour market endogenously separates into submarkets with identical agents on either side of the market.[6] Still, workers' on-the-job search may nonetheless impose externalities on the employer. In models with competitive on-the-job search, it is therefore common to let the contract space be sufficiently rich to align the worker's and firm's incentives. In this case, workers' on-the-job search behaviour is efficient in the sense that it maximizes the joint income of the worker-firm pair, see e.g. Moen and Rosén (2004) and Menzio and Shi (2010). However, as worker income and firm profit is taxed at different rates, it is not obvious how the joint income should be defined. We show that the optimal contract ensures that the worker's on-the-job search maximizes L, the weighted sum of the worker's after-tax utility and the firm's after-tax profit as defined by (3). This require some technical machinery, which we lay out in the paragraph below. Readers without interest in such details may jump directly to the characterization of the competitive search equilibrium in Proposition 1.

[6] Moen (1997) shows that the labour market endogenously separates into submarkets if the workers have different incomes while searching. Searching workers with low current income join submarkets with a high job finding rate and relatively low wages, while the opposite holds for workers with high current income.

Employment contract details. For any $\mathbf{y}^\ell \in \mathscr{Y}$, a submarket opens up. We will later show that submarkets depends only on $\mathbf{y}^\ell$. A submarket is characterized by a contract offered by the firms, $C(\mathbf{y}^\ell)$, and a labour market tightness, $\theta(\mathbf{y}^\ell)$. For workers in the most attractive job, $\theta = 0$. The contract $C(\mathbf{y}^\ell)$ consists of two parts. The first part is a *wage contract*, $C^w(\mathbf{y}^\ell)$. This is a standard component of a competitive search equilibrium. The wage contract specifies a wage function $w(\mathbf{y}, \mathbf{y}^\ell)$, and also specifies which $\mathbf{y}$-draws lead to a job offer; that is, the wage contract includes an attribute acceptance set $\mathscr{Y}_a(\mathbf{y}^\ell) \subseteq \mathscr{Y}$. An applicant is hired if and only $\mathbf{y} \in \mathscr{Y}_a(\mathbf{y}^\ell)$.[7]

The second part of the contract is a *search contract*, C^s, and this component is usually not formalized in models of competitive on-the-job search, but governs job search in the new type-$\mathbf{y}$ job. First, it specifies the worker's search effort in the new job, i.e. $e(\mathbf{y})$. Second, it specifies the worker's job acceptance decision when getting a job offer in the new job. The worker accepts an offer in the new job if and only if the offered NPV utility exceeds a threshold function $\widehat{L}(\mathbf{y})$. Third, it includes a submarket selection function $\sim(\mathbf{y})$, dictating which submarket the worker should search in. The selection function ranks any two submarkets (C, θ) and (C', θ'), in and out of equilibrium, and prescribe that the worker searches in the highest ranked submarket. If two markets have equal rank, i.e. if $(C, \theta) \sim (C', \theta')$, the worker choose freely between them.[8]

The selection function is important in that it pins down out-of-equilibrium beliefs. Suppose that, in equilibrium, all firms that attract workers employed in type-$\mathbf{y}^\ell$ jobs post the contract C and face tightness θ. If a measure-0 set of firms deviate and post C', a new submarket forms. Suppose that the deviating firms still attract workers in type-$\mathbf{y}^\ell$ jobs; then, the selection function ensures that labour market tightness in the new submarket, θ', satisfy $(C, \theta) \sim (C', \theta')$.[9]

Hence, the search contract is $C^s = \Big(e(\mathbf{y}), \widehat{L}(\mathbf{y}), \sim(\mathbf{y})\Big)$ and specifies search behaviour as a function of realized job attributes, $\mathbf{y}$. As will be clear below, the firms do not want to make any aspect of the search contract contingent on the job attributes in the market in which the contract (or vacancy) is posted. Hence, C^s is the same for all firms. It follows that $C(\mathbf{y}^\ell) = (C^w(\mathbf{y}^\ell), C^s)$. The wage and search contracts for given realization of $\mathbf{y}$ are $C^w(\mathbf{y}, \mathbf{y}^\ell)$ and $C^s(\mathbf{y})$, respectively. Finally, unemployed workers search so as to maximize their NPV utility V_0.

Definition 1 (Competitive search equilibrium) *Competitive search equilibrium specifies wage contracts $C^w(\mathbf{y}^\ell)$, a search contract C^s, a labour market tightness $\theta(\mathbf{y}^\ell)$, vacancy returns $\pi(\mathbf{y}^\ell)$, and asset value functions $J(\mathbf{y}, \mathbf{y}^\ell)$ and $V(\mathbf{y}, \mathbf{y}^l)$ such that*

1. *$\big(C^w(\mathbf{y}^\ell), C^s\big)$ maximizes $\pi(\mathbf{y}^\ell)$ given that the workers' search behaviour in the $\mathbf{y}^\ell$ market is governed by $\tilde{C}^s(\mathbf{y}^\ell)$, for all $\mathbf{y}^\ell \in \mathscr{Y}$.*

2. *$\pi(\mathbf{y}^\ell) = \overline{\pi}$ for all $\mathbf{y}^\ell \in \mathscr{Y}$, where $\overline{\pi}$ is defined by (6).*

3. *$V(\mathbf{y}, \mathbf{y}^\ell)$, $J(\mathbf{y}, \mathbf{y}^\ell)$, and $\pi(\mathbf{y}^\ell)$ are given (1), (2), and (5), respectively.*

[7] We also require that the contract prescribed to workers is sufficiently attractive to be accepted when the attribute draw is within the acceptance set. This requirement is trivial to satisfy, and is not spelled out.

[8] We assume that the contract prescribes search behaviour directly. Alternatively, the firm can govern a worker's search by a quit fee. The worker will search to maximize the gain from search, subject to the quit fee. Such a fee is sufficient to ensure both efficient on-the-job search and job acceptances by the worker.

[9] For details on out-of-equilibrium beliefs in competitive search equilibrium, see Moen (1997) and Guerrieri, Shimer, and Wright (2010), and Garibaldi, Moen, and Sommervoll (2016) which deals with on-the-job search.

Before characterizing the equilibrium we show how the optimal search contract is set by the firm as this contract component is non-standard in models of competitive search. For a given wage contract, let $\overline{V}(\mathbf{y}^\ell) = E^{\mathbf{y}|\mathbf{y}\in\mathscr{Y}_a} V(\mathbf{y},\mathbf{y}^\ell)$ be the expected NPV income to a worker if hired by a firm. We refer to this as the promised value to the worker. As in Moen and Rosen (2011), a firm's maximization problem can be divided into two stages.[10] First, for a given promised value $\overline{V}$, and hiring decision $\mathscr{Y}_a$, choose the contract that maximize expected profits $E^{\mathbf{y}|\mathbf{y}\in\mathscr{Y}_a} J(\mathbf{y},\mathbf{y}^\ell)$. Second, choose $\overline{V}$ and $\mathscr{Y}_a$ so as to maximize $\pi(\mathbf{y}^\ell)$. The first stage ensures that the search contract is Pareto efficient. The second stage determines the division of the surplus between the worker and the firm that maximizes the return to the vacancy, as firms optimally trade off a high wage bill and a high arrival rate of job applicants. This trade-off is core in competitive search. The second stage also ensures an efficient hiring decision.

From (3) it follows that

$$E^{\mathbf{y}|\mathbf{y}\in\mathscr{Y}_a} J(\mathbf{y},\mathbf{y}^\ell) \;=\; \frac{1-\tau}{1-t} E^{\mathbf{y}|\mathbf{y}\in\mathscr{Y}_a}\left[L(\mathbf{y}) - V(\mathbf{y},\mathbf{y}^\ell)\right] = \frac{1-\tau}{1-t}\left[E^{\mathbf{y}|\mathbf{y}\in\mathscr{Y}_a} L(\mathbf{y}) - \overline{V}\right] \quad (14)$$

Hence, *the search contract is set so as to maximize the joint tax-adjusted value of a match,* $L(\mathbf{y})$, as defined by (3). This maximization problem is independent of the worker's employment history $\mathbf{y}^\ell$. Thus, $\widehat{L}(\mathbf{y}) = L(\mathbf{y})$, $e(\mathbf{y}) = \arg\max_e L(\mathbf{y})$, and $\sim(\mathbf{y})$ dictates that a worker chooses the submarket that maximizes $L(\mathbf{y})$.

Proposition 1 characterizes the competitive search equilibrium.

Proposition 1 *Let* $\mathbf{y}$ *be the attributes of a worker's current job,* $\mathbf{y}^\ell$ *the attributes of her previous job, and* $\mathbf{y}'$ *the attributes of her new (i.e. next) job. With linear taxes, competitive search equilibrium is determined by the following conditions:*

1. $\mathscr{Y}_a(\mathbf{y}) = \{\mathbf{y}' \in \mathscr{Y} | y_p(1-t) + y_z < y_p'(1-t) + y_z'\}$. *Hence, if* $\mathbf{y}$ *is the type of the current job, and* $\mathbf{y}'$ *the type of the new job, the worker switches job if and only if*

$$y_p(1-t) + y_z < y_p'(1-t) + y_z.' \quad (15)$$

2. *Expected match surplus is split such that* $\overline{V}(\mathbf{y}^\ell) = (1-\beta)L(\mathbf{y}^\ell) + \beta E^{\mathbf{y}|\mathbf{y}\in\mathscr{Y}_a} L(\mathbf{y})$*; hence,*

$$(r+s)L(\mathbf{y}) = y_p(1-t) + y_z - c(e) + sL_0^i + e(\mathbf{y})p(\theta(\mathbf{y}))\beta E^{\mathbf{y}'\in\mathscr{Y}_a(\mathbf{y})}(L(\mathbf{y}') - L(\mathbf{y})) \quad (16)$$

3. *Search effort* $e(\mathbf{y})$ *maximizes* $L(\mathbf{y})$, *with first order conditions given by*

$$c'(e(\mathbf{y})) = p(\theta(\mathbf{y}))\beta E^{\mathbf{y}'\in\mathscr{Y}_a(\mathbf{y})}(L(\mathbf{y}') - L(\mathbf{y})) \quad (17)$$

4. *Labour market tightness* $\theta(\mathbf{y})$ *solves*

$$(1-\beta)q(\theta(\mathbf{y}))E^{\mathbf{y}'\in\mathscr{Y}_a(\mathbf{y})}[L(\mathbf{y}') - L(\mathbf{y})] = \pi\frac{1-t}{1-\tau} \quad (18)$$

[10] Here, and below, the expectation $E^{\mathbf{y}\in\mathscr{Z}} Z(\mathbf{y}) \equiv E^{\mathbf{y}}\mathbf{1}(\mathbf{y}\in\mathscr{Z})Z(\mathbf{y})$ for any function $Z(\mathbf{y})$ and any set $\mathscr{Z}$.

where $\overline{\pi}$ is given by (6)

Proof. See Appendix A. ■

The equilibrium is characterized by the five equations (6) and (15)-(18), and has a familiar structure: the tax adjusted joint income plays the same role as the joint income in standard competitive models of on-the-job search, see Garibaldi, Moen, and Sommervoll (2016).

Competitive search equilibrium pins down $\overline{V}$ (and $\mathscr{Y}_a$), but not wages as a function of the realized $\mathbf{y}'$. That is, the wage paid once a worker is hired has no allocative role, a reflection of the fact that risk neutral workers have preferences only over expected wages. In the empirical part of the paper we assume that the wage sharing rule in the first part of Proposition 1 holds for all $\mathbf{y}$, and hence that

$$V(\mathbf{y}, \mathbf{y}^{\ell}) = \beta L(\mathbf{y}) + (1-\beta)L(\mathbf{y}^{\ell}) \tag{19}$$

for all $\mathbf{y}$. This is clearly consistent with competitive search equilibrium, and is also consistent with a wage contract that specifies a bargaining protocol rather than a wage. Indeed, we show in Appendix B that the competitive search allocation also obtains when wages and search effort are determined by bargaining under the Hosios (1990) efficiency condition.

Lemma 1 highlights some interesting features of the model.

Lemma 1 *The following is true:*

1. *Competitive search equilibrium with labour income tax t is isomorphic to competitive search equilibrium with $t = 0$, with all productivity outcomes y_p scaled down to $(1-t)y_p$, and with $\overline{\pi}$ scaled down to $(1-t)\overline{\pi}$.*

2. *If the vacancy posting costs c^v and the entry costs K are fully deductible ($\gamma_c = \gamma_K = 1$), the profit tax τ does not influence the equilibrium allocation.*

3. *If γ_c and/or γ_K are different from 1, an increase in τ only influences equilibrium through its impact on the vacancy return requirement $\overline{\pi}\frac{1-t}{1-\tau}$.*

The first point follows directly from (16) and (18), as neither L nor the vacancy return requirement will be influenced by the transformation. The second point follows from the fact that with $\gamma_c = \gamma_k = 1$, (6) gives that $\overline{\pi} = (rK + c^v)(1-\tau)$, hence the right-hand side of (18) is equal to $rK + c^v$ and hence independent of τ. The third point follows from the fact that τ only enters the right-hand side of the equilibrium equation (18) in the equilibrium characterization.

The intuition behind the first result is as follows: The zero profit condition implies that income is either allocated to workers in the form of wages, or to the firms to cover their search costs (including the flow equivalent of the entry cost). Labour income tax is levied on the former, but not the latter, making job creation relatively cheaper. Note further that the set of feasible solutions that satisfy the zero profit criterion (18) in the presence of taxes is the same as in the transformed economy with no taxes, but with y_p and $\overline{\pi}$ scaled down with a factor $1-t$. We know that the competitive search allocation without taxes maximizes the NPV joint worker-firm income in all submarkets given the zero profit constraint of firms. From the first result in Lemma 1 it then follows that the equilibrium allocation in the scaled economy maximizes $L(\mathbf{y}^\ell)$:

Corollary 1 *For all $\mathbf{y}^\ell \in \mathscr{Y}$, the equilibrium allocation maximizes $L(\mathbf{y}^\ell)$ given the zero profit constraint of firms.*

In particular, the equilibrium allocation maximizes after-tax utility of unemployed workers, V_0.

As is common in search models with identical firms, equilibrium is defined solely in terms of asset values and labour market tightnesses. Given $\mathscr{Y}_a(\mathbf{y})$, $\theta(\mathbf{y}^\ell)$ and $e(\mathbf{y})$, the steady state stocks of workers in different submarkets are pinned down. See Appendix C for details.

The acceptance set $\mathscr{Y}_a(\mathbf{y})$ in Proposition 1 implies a reservation amenity function:

Definition 2 (The reservation amenity function) *The reservation amenity function*

$$\phi(y_p', \mathbf{y}) \equiv y_z - (1-t)(y_p' - y_p)$$

returns the minimum level of amenity required for a worker currently employed in a job with attributes $\mathbf{y}$ to accept employment at job with productivity attribute y_p'.

Clearly, $\mathbf{y}' \in \mathscr{Y}_a(\mathbf{y})$ if and only if $y_z' \geq \phi(y_p', \mathbf{y})$.

4 Optimal Redistributive Taxation

We now introduce a social planner who puts welfare weights on different types of workers depending on their expected NPV utility (in unemployment), $V_0^i + t_0$, where t_0 is a lump sum transfer. In doing so we follow Golosov, Maziero, and Menzio (2013), Best and Kleven (2013) and others.[11] The welfare function is

$$\Omega = \sum_{i=1}^{I} \Phi(V_0^i + t_0), \tag{20}$$

where Φ is a strictly increasing and concave function.

Suppose the planner in steady state needs to raise an amount G in NPV income. The planner takes job attributes $\mathbf{y}$ and the equilibrium responses of workers and firms as given, including their ranking of jobs. The planner's objective concerns the NPV's of the unemployed workers. As argued in Pissarides (2000, p. 187), the NPV utility of an unemployed worker is maximized in the efficient solution that maximizes total output in the economy. We assume that the planner takes a long-term view and is concerned only about the ergodic state of the economy when

[11] The introduction of a risk averse planner into an economy populated by risk neutral workers may be justified if, in the background, large families of identical risk-averse workers pool idiosyncratic income risk.

designing the tax system.[12] For now we assume that the investment cost K and the search cost of firms c^v are fully deductible, which is equivalent to setting $\tau = 0$, see (6). The planner's problem can therefore be expressed as

$$\max_{t_0,t} \sum_{i=1}^{I} \kappa_i \Phi(V_0^i + t_0) \quad \text{subject to} \quad \sum_{i=1}^{I} \kappa_i (tW_0^i - B_0^i - G - t_0) \geq 0, \tag{21}$$

where the inequality is the planner's budget constraint, and B_0^i is the NPV of the gross income flow from unemployed type-i workers, see (9).

Let λ denote the Lagrange multiplier associated with the constraint, interpretable as the planner's valuation of a marginally increased budget. From (1) and (12) we have that $V_0^i = Y_{z,0}^i + (1-t)W_0^i$. The Envelope Theorem implies that $\partial V_0^i/\partial t = -W_0^i$. It follows that the first order conditions for the planner's problem read

$$\sum_{i=1}^{I} \kappa_i \Phi'(V_0^i + t_0) = \lambda, \tag{22}$$

$$\sum_{i=1}^{I} \kappa_i \Phi'(V_0^i + t_0) W_0^i = \lambda \sum_{i=1}^{I} \kappa_i \left(W_0^i + t\frac{\partial W_0^i}{\partial t} - \frac{\partial B_0^i}{\partial t} \right), \tag{23}$$

$$\sum_{i=1}^{I} \kappa_i t(W_0^i - B_0^i) - G - t_0 = 0. \tag{24}$$

Equation (22) shows that the planner values additional tax revenue by the welfare gains it yield, averaged across worker types. Let $\overline{W}_0 \equiv \sum_{i=1}^{I} \kappa_i W_0^i$ be the tax base in the economy, and let $\overline{B}_0 \equiv \sum_{i=1}^{I} \kappa_i B_0^i$ be the NPV of the gross income flow from the unemployed workers. Then, (22) and (23) implies

$$-\operatorname{Cov}\big(\Phi'(V_0^i + t_0), W_0^i\big) = \lambda \left(-t\frac{\partial \overline{W}_0}{\partial t} + \frac{\partial \overline{B}_0}{\partial t} \right), \tag{25}$$

or alternatively,

$$-\frac{\operatorname{Cov}\big(\Phi'(V_0^i + t_0), W_0^i\big)}{\lambda \overline{W}_0} = \mathrm{El}_t \overline{W}_0 + \frac{\partial \overline{B}_0/\partial t}{\overline{W}_0} \tag{26}$$

where $\mathrm{El}_t \overline{W}_0 = \left| \frac{t}{\overline{W}_0} \frac{\partial \overline{W}_0}{\partial t} \right|$ denotes the (positive) elasticity of the tax base W_0 with respect to t.[13]

The left-hand side of (26) represents the equity gain to the planner of increasing t and redistributing the tax revenue to all the agents through higher transfers, t_0. The equity gain is normalized by the total tax base measured in terms of the planner's utility of public funds. Note that the left-hand side of (26) is close to zero when t is close to one and otherwise strictly

[12] Suppose at time 0, all workers enter the economy as unemployed. At that point the planner sets the tax rates. Since the asset value equations will be constant through time, the planner will never want to change the tax rates.

[13] Equation (26) obtains by rewriting (23) as $\sum_{i=1}^{I} \kappa_i \Phi'(V_0^i + t_0) W_0^i - \lambda \overline{W}_0 = \lambda \left(t\frac{\partial W_0^i}{\partial t} + B_0^i - (1-t)\frac{\partial B_0^i}{\partial t} \right)$. Equation (22) now implies the left-hand side is $\operatorname{Cov}\left(\Phi'(V_0^i + t_0), W_0^i\right)$, and (26) follows by application of the El_t-operator, noting that $\partial \overline{W}_0/\partial t \leq 0$.

positive. The right-hand side of (26) is the sum of the elasticity of the tax base $\overline{W}_0$, and the increase in the aggregate unemployment benefit expenditures relative to the tax base, and reflects the deadweight loss of income taxation. The deadweight loss of taxation thus operates both through a revenue channel, captured by $\mathrm{El}_t\overline{W}_0$, and an expenditure channel, captured by $\frac{\partial\overline{B}_0/\partial t}{\overline{W}_0}$. The deadweight loss is zero when t and $\overline{B}_0$ are both zero.

Equity gains are larger when the planner has strong equity concerns, such that $\Phi''(\cdot)$ is strongly negative, and when $\Phi'(V_0^i + t_0)$ and W_0^i have high negative covariance. Because $\Phi(\cdot)$ is strictly concave, $\Phi'(V_0^i + t_0)$ and W_0^i have high negative covariance when V_0^i and W_0^i have high positive covariance, i.e. when the utility enjoyed by a particular worker-type is highly correlated with that worker-type's tax base. Redistributive taxation is particularly attractive to the planner in this case because income taxation targets the high-utility worker-types.[14] According to (26), the optimal tax rate t exactly balances the equity gain and the deadweight loss associated with a marginal tax increase.

5 Deadweight Losses

The tax and transfer system gives rise to deadweight losses along three margins: job search effort $e(\mathbf{y})$, job ranking, represented by the reservation amenity function $\phi(\cdot,\mathbf{y})$, and vacancy creation, $\theta(\mathbf{y})$. We analyse spell-specific partial marginal deadweight losses: the deadweight losses that arise from marginal distortions to one of the three endogenous variables, $e(\mathbf{y})$, $\phi(\cdot,\mathbf{y})$, and $\theta(\mathbf{y})$, in a particular spell at a particular job ladder rung-$\mathbf{y}$, *holding the values of the other endogenous variables (and thus the tax bases) at all future job or unemployment spells constant.* The spell-specific partial deadweight losses are analytically tractable (details in Appendix D), fully encode the economic mechanisms in play, and integrates to the total marginal deadweight loss.

5.1 Preliminaries

It is convenient to derive the effect of taxation on M_0^i, and subsequently back out the desired effects on W_0^i.[15] Since $M_0^i \equiv V_0^i + tW_0^i - B_0^i$ under the maintained assumption that investment and firm search costs K and v are fully deductible, and because search effort, job rankings, and vacancy posting in the unemployment-submarket maximizes $L_0^i \equiv V_0^i = Y_{0,z}^i + (1-t)W_0^i$, the Envelope Theorem implies $\partial V_0^i/\partial t = -W_0^i$, and

$$-\frac{\partial M_0^i}{\partial t} = -t\frac{\partial W_0^i}{\partial t} + \frac{\partial B_0^i}{\partial t}. \tag{27}$$

This is the deadweight loss for type-i workers. The left-hand side shows the marginal reduction in total NPV utility created by unemployed type-i workers when t increases marginally. The

[14] Worker-type utility V_0^i and worker-type tax bases W_0^i are not necessarily aligned in our model due to the presence of amenities; indeed, $V_0^i = Y_{z,0}^i + (1-t)W_0^i$.

[15] The income tax distortions to the tax base W_0 depends on the distortions to all the submarkets an unemployed worker may subsequently find herself in, i.e. on distortions at every rung of the job ladder. Since the wage of particular worker at a particularly job ladder rung depends on both the type-$\mathbf{y}$ of the current job and the type-$\mathbf{y}^\ell$ of the the previous job, it is cumbersome to work directly with W_0.

right-hand side shows that the deadweight loss arises from two fiscal externalities: a reduction in the government's tax revenue from worker type-i and a change in the unemployment benefit expenditures on worker type-i. The total deadweight loss on right-hand side of (26) obtains by aggregating (27) across worker types.

Remark 1 *$M(\mathbf{y})$ can be represented as $M(\mathbf{y}) = \tilde{M}(e(\mathbf{y}), \phi(\cdot, \mathbf{y}), \theta(\mathbf{y}), M(\cdot))$, where $e(\mathbf{y})$, $\phi(\cdot, \mathbf{y})$ and $\theta(\mathbf{y})$ are search effort, reservation amenity function, and labour market tightness in the current rung-$\mathbf{y}$ spell, and $\tilde{M}$ is the functional implicitly defined by (10).*

Ordinarily, interest centers on the properties, i.e. existence and uniqueness, of a fixed point of $M(\mathbf{y}) = \tilde{M}(e(\mathbf{y}), \phi(\cdot, \mathbf{y}), \theta(\mathbf{y}), M(\cdot))$. Here, however, we exploit that the functional $\tilde{M}$ allows us to define *single-spell partial derivatives* of $M(\mathbf{y})$ with respect to job search, reservation amenities and tightness at rung-$\mathbf{y}$.

Definition 3 (Single-spell partial derivative) *The single-spell partial derivative of $M(\mathbf{y})$ with respect to $x(\mathbf{y}) \in \{e(\mathbf{y}), \phi(y_p', \mathbf{y}), \theta(\mathbf{y})\}$ is the partial derivative of $\tilde{M}$ with respect to $x(\mathbf{y})$, holding the continuation value function $M(\cdot)$ constant. We use the shorthand*

$$\frac{\partial \tilde{M}(\mathbf{y})}{\partial x(\mathbf{y})} \equiv \frac{\partial}{\partial x(\mathbf{y})} \tilde{M}(e(\mathbf{y}), \phi(\cdot, \mathbf{y}), \theta(\mathbf{y}), M(\cdot)),$$

for the single-spell partial derivative of $M(\mathbf{y})$ with respect to $x(\mathbf{y})$.

$\frac{\partial \tilde{M}(\mathbf{y})}{\partial x(\mathbf{y})}$ has a straightforward interpretation as the marginal effect on $M(\mathbf{y})$ of a distortion to $x(\mathbf{y}) \in \{e(\mathbf{y}), \phi(y_p', \mathbf{y}), \theta(\mathbf{y})\}$ in the *current* rung-$\mathbf{y}$ spell only, i.e. holding NPV values $M(\cdot)$ in all future job or unemployment spells constant.

With a continuous set of job attributes $\mathscr{Y} \setminus \{\mathbf{y}_0\}$, a worker returning to the same job ladder rung is a zero probability event, and the stipulation that $\frac{\partial \tilde{M}(\mathbf{y})}{\partial x(\mathbf{y})}$ holds the $M(\mathbf{y})$-value in future rung-$\mathbf{y}$ job spells constant is immaterial. However, because workers always return to unemployment following a job destruction shock, the stipulation does have bite for the measurement of partial effects of distortions in the unemployment submarkets. We emphasize therefore that, for $\mathbf{y} = \mathbf{y}_0$, Definition 3 means that $\frac{\partial \tilde{M}(\mathbf{y}_0)}{\partial x(\mathbf{y}_0)}$ is the partial effect on $M(\mathbf{y}_0)$ of a distortion to $x(\mathbf{y}_0)$ in the *current* unemployment spell only. This ensures consistency of single-spell partial derivatives in employment and unemployment submarkets.

5.2 Single-Spell Partial Deadweight Losses in Employment

Consider any employment rung $\mathbf{y} \in \mathscr{Y} \setminus \{\mathbf{y}_0\}$ on the job ladder. Define

$$R(\mathbf{y}) \equiv \left[r + s + e(\mathbf{y})p(\theta(\mathbf{y})) \Pr(\mathbf{y}' \in \mathscr{Y}_a(\mathbf{y}))\right]^{-1},$$

where $\Pr(\mathbf{y}' \in \mathscr{Y}_a(\mathbf{y}))$ is the probability that the worker quits after meeting a vacancy.[16]

[16] $R(\mathbf{y})$ is the present discounted value of a stream of unit payments terminated at rate $s + p(\theta(\mathbf{y}))e(\mathbf{y}) \Pr(\mathbf{y}' \in \mathscr{Y}_a(\mathbf{y}))$. Indeed, if $\eta \equiv s + p(\theta(\mathbf{y}))e(\mathbf{y}) \Pr(\mathbf{y}' \in \mathscr{Y}_a(\mathbf{y}))$, then $\int_0^\infty \left[\int_0^t e^{-rs} ds\right] \eta e^{-\eta t} dt = [r + \eta]^{-1}$.

Job search effort. We show in Appendix D that the incremental distortion to job search effort in a rung-$\mathbf{y}$ job spell following a marginal increase in the tax rate t is given by

$$\frac{\partial e(\mathbf{y})}{\partial t} = -\frac{p(\theta(\mathbf{y}))}{c''(e(\mathbf{y}))}\left[\frac{\beta(1-t)\Delta Y_p(\mathbf{y}) - (1-\beta)\Delta Y_z(\mathbf{y})}{1-t}\right]. \tag{28}$$

The direction of the distortion depends on whether on-the-job search effort, on the margin, gives rise to higher or lower expected wage income. Search effort is associated with higher expected wages when $\Delta Y_p(\mathbf{y})$ is large relative to $\Delta Y_z(\mathbf{y})$, i.e. when a submarket-$\mathbf{y}$ job promises lots of scope for productivity growth relative to amenity growth. In that case, higher income taxation disincentivises search effort, and $\frac{\partial e(\mathbf{y})}{\partial t} < 0$. When $\Delta Y_p(\mathbf{y})$ is small relative to $\Delta Y_z(\mathbf{y})$, higher income taxation in fact incentivises search effort, and $\frac{\partial e(\mathbf{y})}{\partial t} > 0$.

Distortions to workers' on-the-job search effort in a submarket-$\mathbf{y}$ job spell has allocative effects and generates a single-spell partial deadweight loss $O_t^e(\mathbf{y}) \equiv -\frac{\partial \tilde{M}(\mathbf{y})}{\partial e(\mathbf{y})}\frac{\partial e(\mathbf{y})}{\partial t}$ in the form of a reduced rung-$\mathbf{y}$ tax base. We show in Appendix D that

$$O_t^e(\mathbf{y}) = t\frac{R(\mathbf{y})p(\theta(\mathbf{y}))^2}{c''(e(\mathbf{y}))}\left[\frac{\beta(1-t)\Delta Y_p(\mathbf{y}) - (1-\beta)\Delta Y_z(\mathbf{y})}{1-t}\right]^2, \tag{29}$$

which is strictly positive for almost all $\mathbf{y} \in \mathscr{Y} \setminus \{\mathbf{y}_0\}$: In response to a marginal increase in t, on-the-job search effort in a rung-$\mathbf{y}$ spell falls when the rung-$\mathbf{y}$ tax base is increasing in search effort, and increases when the rung-$\mathbf{y}$ tax base is decreasing in search effort. In effect, distortions to search effort give deadweight losses because they shift employment away from high-productive, low-amenity (high-tax base) jobs, towards low-productive, high-amenity (low-tax base) jobs.

Job ranking. A marginal change in the tax rate impacts the submarket-$\mathbf{y}$ reservation amenity $\phi(y_p', \mathbf{y})$ associated with a productivity attribute draw y_p'. From Definition 2:

$$\frac{\partial \phi(y_p', \mathbf{y})}{\partial t} = y_p' - y_p; \tag{30}$$

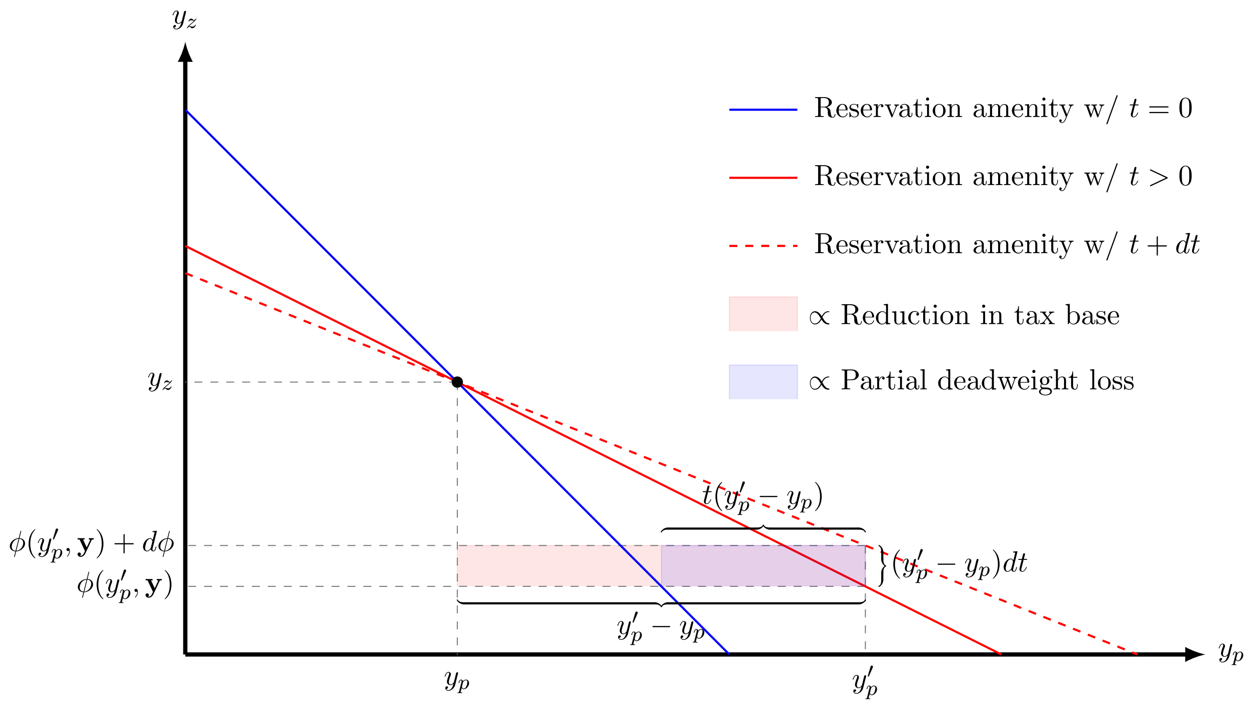
that is, a higher income tax increases (decreases) the reservation amenity for alternative type-$\mathbf{y}'$ job with higher (lower) productivity attribute than the current type-$\mathbf{y}$ job. A higher income tax makes high-productivity jobs less attractive and low-productivity jobs more attractive.

A distorted ranking of jobs impact the direction of worker flows. These distortions have no first order effects on worker utility, but their effects on the employment allocation does give rise to deadweight losses. Indeed, the single-spell partial deadweight loss from distorted amenity reservation for productivity draw y_p' in a rung-$\mathbf{y}$ job spell is $-\frac{\partial \tilde{M}(\mathbf{y})}{\partial \phi(y_p', \mathbf{y})}\frac{\partial \phi(y_p', \mathbf{y})}{\partial t}$. We show in Appendix D that, when integrated across productivity attribute draws y_p', the single-spell partial deadweight loss from distorted amenity reservation function is

$$O_t^\phi(\mathbf{y}) = tR(\mathbf{y})^2 e(\mathbf{y})p(\theta(\mathbf{y}))\int (y_p' - y_p)^2 f_p(y_p') f_z(\phi(y_p', \mathbf{y}))dy_p', \tag{31}$$

which is strictly positive. On the margin, a higher tax rate steers workers away from high-productive, low-amenity jobs with larger tax bases and towards low-productive, high-amenity jobs with smaller tax bases.

Figure 2: Tax Distortion to Job Rankings



Notes: The figure illustrates the implications, in a type-$\mathbf{y}$ submarket, of a marginal tax increase, $dt > 0$ for the reservation amenity associated with a job offer with productivity attribute $y_p' > y_p$. As detailed in the text, an increase in the tax rate also entails a deadweight loss in in the case where $y_p' < y_p$.

Figure 2 graphs the reservation amenity levels $\phi(y_p', \mathbf{y})$ for different tax rates, $t = 0$, $t > 0$, and $t + dt > t$, for a worker who is employed in a job with attributes $\mathbf{y}$; of course, the worker accepts attribute combinations above the reservation amenity graph. Following an increase in the tax rate from t to $t + dt$, the reservation amenity $\phi(y_p', \mathbf{y})$ increases to $\phi(y_p', \mathbf{y}) + d\phi = \phi(y_p', \mathbf{y}) + (y_p' - y_p)dt$. This reduces the tax base by $(y_p' - y_p)^2 dt$, represented by the red-shaded rectangle in Figure 2, and gives rise to a deadweight loss $t(y_p' - y_p)^2 dt$, represented by the blue-shaded rectangle with area (assuming a tax rate of 50 percent for illustration).[17]

Vacancy creation. Firms can provide value to workers through high wages upon meeting a vacancy or through high job finding rates (i.e. higher labour market tightness). In competitive search equilibrium, promised remuneration and labour market tightness is balanced to maximize workers' after-tax utility subject to firms' profit requirement under free entry. Income taxation may distort this balance away from the efficient benchmark. We show in Appendix D that

$$\frac{\partial \theta(\mathbf{y})}{\partial t} = \frac{\theta(\mathbf{y})\Delta Y_z(\mathbf{y})}{\beta(1-t)\Delta L(\mathbf{y})} = \frac{\theta(\mathbf{y})\Delta Y_z(\mathbf{y})}{\beta(1-t)\left[(1-t)\Delta Y_p(\mathbf{y}) + \Delta Y_z(\mathbf{y})\right]}. \tag{32}$$

If $\Delta Y_z(\mathbf{y}) > 0$, submarket-$\mathbf{y}$ workers expect to improve their position on the amenity ladder. As amenities are untaxed, an increase in the tax rate tilts workers' preference in favour of a higher

[17]The expression for $O_t^{\phi(y_p', \mathbf{y})}(\mathbf{y})$ in (31) obtains by multiplying by $R(\mathbf{y})^2 e(\mathbf{y}) p(\theta(\mathbf{y}))$ and integrating over y_p'.

job finding rate, whereas firms' trade-offs are unaffected, and competitive search equilibrium therefore prescribes an increase in labour market tightness, $\frac{\partial \theta(\mathbf{y})}{\partial t} > 0$. If $\Delta Y_z(\mathbf{y}) < 0$, workers require higher promised remuneration to search from submarket-$\mathbf{y}$; hence, $\frac{\partial \theta(\mathbf{y})}{\partial t} < 0$

Interestingly, if $\Delta Y_z(\mathbf{y}) = 0$, a tax increase has no distortionary effect on vacancy creation. Without an amenity ladder, higher expected wages and a higher job finding rate are both purely pecuniary gains that are taxed at the same proportional rate. This leaves workers' preferences for the two forms of value creation, and hence their equilibrium provision, unaffected.

The rung-$\mathbf{y}$ single-spell partial deadweight loss from distorted rung-$\mathbf{y}$ vacancy creation is $O_t^\theta(\mathbf{y}) \equiv -\frac{\partial \tilde{M}(\mathbf{y})}{\partial \theta(\mathbf{y})} \frac{\partial \theta(\mathbf{y})}{\partial t}$. We show in Appendix D that

$$O_t^\theta(\mathbf{y}) = tR(\mathbf{y}) \frac{(1-\beta)e(\mathbf{y})p(\theta(\mathbf{y}))}{\beta\left[(1-t)\Delta Y_p(\mathbf{y}) + \Delta Y_z(\mathbf{y})\right]} \left[\frac{\Delta Y_z(\mathbf{y})}{1-t}\right]^2, \tag{33}$$

which is strictly positive if $\Delta Y_z(\mathbf{y}) \neq 0$, but zero if $\Delta Y_z(\mathbf{y}) = 0$ for the reasons given above. When $\Delta Y_z(\mathbf{y}) \neq 0$, distorted vacancy creation gives deadweight losses because it induces workers to leave high-productivity, low-amenity jobs too fast and to remain for too long in low-productivity, high-amenity jobs. As workers are indifferent between these jobs, the shift in the employment allocation has no first-order effects on worker utility, but it does reduce the tax bases.

Figure 3, with submarket-$\mathbf{y}$ tightness $\theta(\mathbf{y})$ on the horizontal axis and the expected before-tax NPV wage after successful on-the-job search, $W^a(\mathbf{y}) = E^{\mathbf{y}' \in \mathscr{Y}^a(\mathbf{y})} W(\mathbf{y}', \mathbf{y})$ on the vertical axis, illustrates the economic forces that shape the deadweight loss from distorted vacancy creation when rung-$\mathbf{y}$ workers expect positive amenity growth, i.e. $\Delta Y_z(\mathbf{y}) > 0$. The zero profit condition (in blue) is represented by a concave iso-profit curve, and workers' indifference curves (in red) are convex to the origin.[18] In Figure 3, firms enjoy higher profits closer to the origin, while workers are better off farther from the origin. For a given tax rate t, equilibrium-(θ, W^a) is the point of tangency between the zero profit condition and the worker indifference curve.[19] A small dt-increase in the income tax rate leaves the zero profit curve unaffected, but shifts workers' trade-off in favour of a higher job finding rate, i.e. renders the indifference curves steeper, as indicated in Figure 3. Hence, equilibrium tightness increases marginally while expected accepted wages falls: workers leave the high-productivity, low-amenity rung-$\mathbf{y}$ submarket at a faster rate.[20]

5.3 Single-Spell Partial Deadweight Losses in Unemployment

Consider now the unemployment rung $\mathbf{y}_0$ of the job ladder. Define

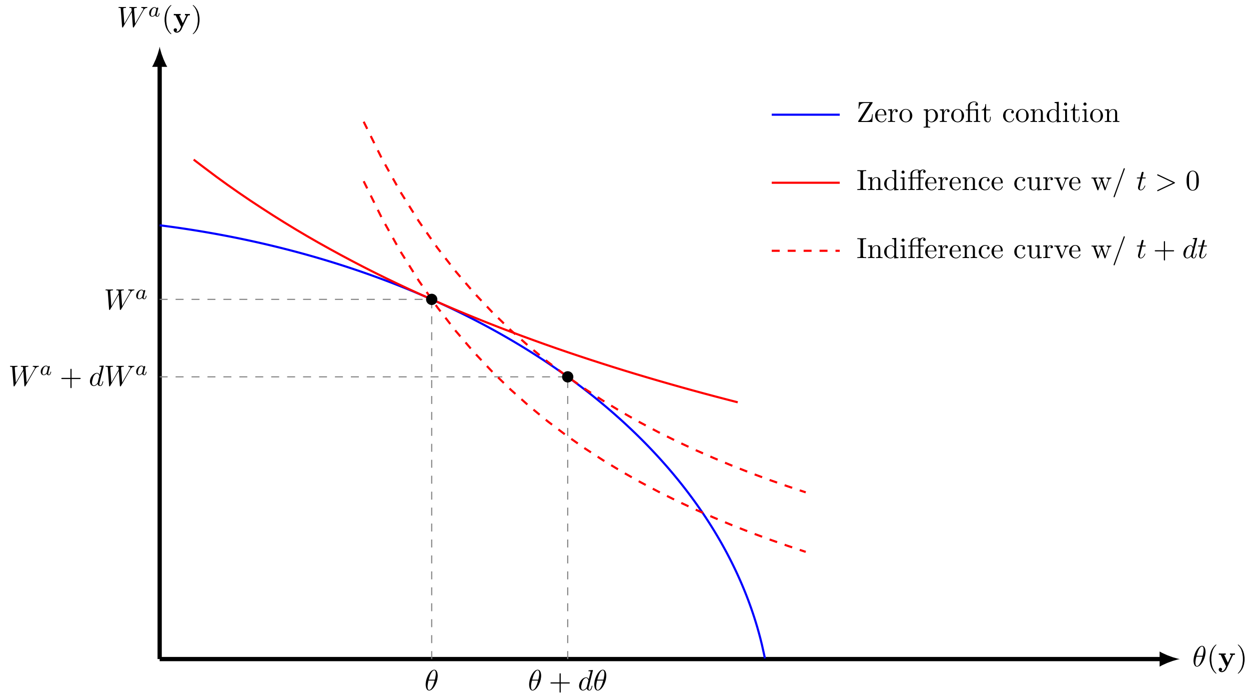
$$R_0 \equiv \left[r + e(\mathbf{y}_0)p(\theta(\mathbf{y}_0))\right]^{-1}.$$

[18]The zero profit curve at rung-$\mathbf{y}$ is implicitly defined by $(1-t)\overline{\pi} = E^{\mathbf{y}' \in \mathscr{Y}^a(\mathbf{y})}\left[Y_p(\mathbf{y}') - W(\mathbf{y}', \mathbf{y})\right]$. Since $E^{\mathbf{y}' \in \mathscr{Y}^a(\mathbf{y})} Y_p(\mathbf{y}')$ is independent of the promised NPV wage $W^a(\mathbf{y})$, the equation defines a unique, decreasing relationship between $W^a(\mathbf{y})$ and $\theta(\mathbf{y})$.

[19]With a Cobb-Douglas matching function, it follows readily that the maximization problem is concave, and the tangency point is unique, see Moen, 1997.

[20]Graphically, when $\Delta Y_z(\mathbf{y}) < 0$, an increase in the income tax rate flattens workers' $(\theta(\mathbf{y}), W^a(\mathbf{y}))$-indifference curves, equilibrium tightness shifts to the left in Figure 3, and workers remain in the low-productivity, high-amenity submarket-$\mathbf{y}$ for longer.

Figure 3: Tax Distortion to Vacancy Creation



Notes: The vertical axis measures $W^a(\mathbf{y}) = E^{\mathbf{y}' \in \mathscr{Y}^a} W(\mathbf{y}', \mathbf{y})$, the average wage paid to workers meeting a vacancy and accepting a job offer in submarket-$\mathbf{y}$. The figure illustrates the implications of a marginal tax increase, $dt > 0$, in a type-$\mathbf{y}$ submarket where $\Delta Y_z(\mathbf{y}) > 0$.

R_0 reflects that unemployed workers are not subjected job destruction shocks and accept all job offers irrespective of the tax rate. There are therefore no deadweight losses from job ranking in the unemployment submarkets, i.e. $O_t^\phi(\mathbf{y}_0) = 0$.

Job search effort. The tax-induced distortion of unemployed job search effort, i.e. $\frac{\partial e(\mathbf{y}_0)}{\partial t}$, obtains by evaluating (28) in $\mathbf{y} = \mathbf{y}_0$. Distorted unemployed job search effort generate a single-spell partial deadweight loss through tax base shrinkage. This deadweight loss component, originating on the revenue side of the government budget, is obtained by replacing $R(\mathbf{y})$ by R_0 in (29) and evaluating the resulting expression at $\mathbf{y} = \mathbf{y}_0$.

Next consider deadweight losses on the expenditure-side of the government budget. The NPV of unemployment benefit provision to an unemployed worker is B_0, see (9), and is $B_1 = \frac{s}{r+s} B_0$ to an employed worker. Distortions to unemployed job search effort impact B_0, which adds to the single-spell partial deadweight loss from job search in the unemployment submarket. Indeed, we show in Appendix D that

$$\frac{\partial \tilde{B}_0}{\partial e(\mathbf{y}_0)} \frac{\partial e(\mathbf{y}_0)}{\partial t} = R_0 p(\theta(\mathbf{y}_0))(B_1 - B_0) \frac{\partial e(\mathbf{y}_0)}{\partial t}, \tag{34}$$

which is strictly positive, i.e. adding to the overall deadweight loss, when $\frac{\partial e(\mathbf{y}_0)}{\partial t} < 0$, reflecting a positive externality of unemployed job search on government expenditures.

Altogether, the single-spell partial deadweight loss from distorted unemployed job search effort, $O_t^e(\mathbf{y}_0)$, is given by (details in Appendix D)

$$O_t^e(\mathbf{y}_0) = -R_0 p(\theta(\mathbf{y}_0)) \left[t\frac{\beta(1-t)\Delta Y_p(\mathbf{y}_0) - (1-\beta)\Delta Y_z(\mathbf{y}_0)}{1-t} - (B_1 - B_0) \right] \frac{\partial e(\mathbf{y}_0)}{\partial t}, \tag{35}$$

which, as long as $\frac{\partial e(\mathbf{y}_0)}{\partial t} < 0$, is strictly positive.

Vacancy creation. Mirroring the distortions to unemployed job search effort, the marginal distortion to vacancy creation from income taxation in the unemployment submarkets, $\frac{\partial \theta(\mathbf{y}_0)}{\partial t}$, obtains by evaluating (32) in $\mathbf{y} = \mathbf{y}_0$. Similarly, the single-spell partial deadweight loss from vacancy creation in the unemployment submarket that arises because of a reduced tax base obtains by replacing $R(\mathbf{y})$ by R_0 in (33) and evaluating at $\mathbf{y} = \mathbf{y}_0$.

Distortions to vacancy creation in the unemployment submarket also impact the expenditure side of the government budget, via B_0. In fact, we show in Appendix D that

$$\frac{\partial \tilde{B}_0}{\partial \theta(\mathbf{y}_0)} \frac{\partial \theta(\mathbf{y}_0)}{\partial t} = R_0 e(\mathbf{y}_0) p'(\theta(\mathbf{y}))(B_1 - B_0) \frac{\partial \theta(\mathbf{y}_0)}{\partial t}, \tag{36}$$

which is negative because $\frac{\partial \theta(\mathbf{y}_0)}{\partial t} > 0$, here reflecting a positive externality of vacancy creation in the unemployment submarket on government expenditures.

Hence, the single-spell partial deadweight loss from vacancy creation in the unemployment submarket $O_t^\theta(\mathbf{y}_0)$ is (details in Appendix D)

$$O_t^\theta(\mathbf{y}_0) = R_0 e(\mathbf{y}_0) p'(\theta(\mathbf{y}_0)) \left[\frac{\Delta Y_z(\mathbf{y}_0)}{1-t} + (B_1 - B_0) \right] \frac{\partial \theta(\mathbf{y}_0)}{\partial t}. \tag{37}$$

Note that, in unemployment, it is necessarily the case that $\Delta Y_z(\mathbf{y}_0) > 0$, such that $\frac{\partial \theta(\mathbf{y}_0)}{\partial t} > 0$. Hence, with respect to job creation, the fiscal externality on the expenditure side of the government budget always counteracts the fiscal externality on the revenue side.

5.4 Aggregation

Workers' careers are sequences of employment cycles during which workers ascend the job ladder, each employment cycle ended by a job destruction event that relocate the worker to rung-$\mathbf{y}_0$, from which a new cycle is initiated when the worker finds a job. Each spell at each job ladder rung in each employment cycle yields single-spell partial deadweight losses of job search, job ranking, and vacancy creation, namely $O_t^e(\mathbf{y})$, $O_t^\phi(\cdot, \mathbf{y})$, and $O_t^\theta(\mathbf{y})$ for $\mathbf{y} \in \mathscr{Y}$. The total marginal deadweight loss obtains by integrating the single-spell partial deadweight losses.

The total marginal deadweight loss is $-\frac{\partial \overline{M}_0}{\partial t}$, where $\overline{M}_0 = \sum_{i=1}^I \kappa_i M_0^i$, see (27). Clearly,

$$-\frac{\partial \overline{M}(\mathbf{y}_0)}{\partial t} = \overline{M}_t^e(\mathbf{y}_0) + \overline{M}_t^\phi(\mathbf{y}_0) + \overline{M}_t^\theta(\mathbf{y}_0), \tag{38}$$

which decomposes the marginal deadweight loss into components stemming from distorted job search effort, $\overline{M}_t^e(\mathbf{y}_0)$, job rankings, $\overline{M}_t^\phi(\mathbf{y}_0)$, and vacancy creation, $\overline{M}_t^\theta(\mathbf{y}_0)$; indeed,

$$\overline{M}_t^e(\mathbf{y}_0) \equiv -\sum_{i=1}^{I} \kappa_i \int_{\underline{y}_p}^{\overline{y}_p} \int_{\underline{y}_z}^{\overline{y}_z} \frac{\partial M^i(\mathbf{y}_0)}{\partial e(\mathbf{y}')} \frac{\partial e(\mathbf{y}')}{\partial t} d\mathbf{y}',$$
$$\overline{M}_t^\phi(\mathbf{y}_0) \equiv -\sum_{i=1}^{I} \kappa_i \int_{\underline{y}_p}^{\overline{y}_p} \int_{\underline{y}_z}^{\overline{y}_z} \int_{\underline{y}_p}^{\overline{y}_p} \frac{\partial M(\mathbf{y}_0)}{\partial \phi(y_p'', \mathbf{y}')} \frac{\partial \phi(y_p'', \mathbf{y}')}{\partial t} dy_p'' d\mathbf{y}',$$
$$\overline{M}_t^\theta(\mathbf{y}_0) \equiv -\sum_{i=1}^{I} \kappa_i \int_{\underline{y}_p}^{\overline{y}_p} \int_{\underline{y}_z}^{\overline{y}_z} \frac{\partial M^i(\mathbf{y}_0)}{\partial \theta(\mathbf{y}')} \frac{\partial \theta(\mathbf{y}')}{\partial t} d\mathbf{y}',$$

where $\underline{y}_p$ and $\overline{y}_p$ are the infimum and supremum of the support of y_p, and $\underline{y}_z$ and $\overline{y}_z$ are the infimum and supremum of the support of y_z.

We show in Appendix E that there exists weights ω_0^i and $\omega_1^i(\mathbf{y})$ such that, for $x \in \{e, \phi, \theta\}$,

$$\overline{M}_t^x(\mathbf{y}_0) = \sum_{i=1}^{I} \kappa_i \left\{ \omega_0^i O_t^{x,i}(\mathbf{y}_0) + \int_{\underline{y}_p}^{\overline{y}_p} \int_{\underline{y}_z}^{\overline{y}_z} \omega_1^i(\mathbf{y}') O_t^{x,i}(\mathbf{y}') d\mathbf{y}' \right\}. \tag{39}$$

Indeed, $\omega_0^i = \frac{r+s}{r}\left[\frac{r+e^i(\mathbf{y}_0)p(\theta^i(\mathbf{y}_0))}{r+s+e^i(\mathbf{y}_0)p(\theta^i(\mathbf{y}_0))}\right]$ and $\omega_1^i(\mathbf{y}') = \omega_0^i \int_{\underline{y}_p}^{\overline{y}_p} \int_{\underline{y}_z}^{\overline{y}_z} \xi^i(\mathbf{y}_0, \mathbf{y}) \xi^{*,i}(\mathbf{y}, \mathbf{y}') d\mathbf{y}$, where $\xi^i(\mathbf{y}_0, \mathbf{y})$ is the discounted density that an unemployment spell ends with the worker moving to a rung-$\mathbf{y}$ job,[21] and $\xi^*(\mathbf{y}, \mathbf{y}')$ is a weight function with an intuitive interpretation that we discuss further in Appendix E; here, it suffices to note that $\xi^*(\mathbf{y}, \mathbf{y}')$ incorporates the likelihood that a sequence of jobs starting at rung-$\mathbf{y}$ and uninterrupted by unemployment involves a spell at rung-$\mathbf{y}'$.

That is, each of $\overline{M}_t^e(\mathbf{y}_0)$, $\overline{M}_t^\phi(\mathbf{y}_0)$ and $\overline{M}_t^\theta(\mathbf{y}_0)$, measuring the deadweight loss from distortions to job search effort, reservation amenities, and vacancy creation, respectively, can be expressed as integrals of the respective single-spell partial marginal deadweight loss functions, $O_t^{e,i}$, $O_t^{\phi,i}$, and $O_t^{\theta,i}$ along the entire job ladder. We use equation (38) with $\overline{M}_t^e(\mathbf{y}_0)$, $\overline{M}_t^\phi(\mathbf{y}_0)$ and $\overline{M}_t^\theta(\mathbf{y}_0)$ given by (39) to compute and decompose the deadweight loss of taxation.

6 Data

Our empirical analysis uses administrative matched employer-employee data covering the entire Danish population during 1994-2003, a period in which Denmark had a stable income tax system. On the worker side, the data contains individual labour market histories measured at a daily frequency, job-specific annual average hourly wages, detailed information on individual tax filings, and a host of socio-economic background characteristics. On the firm side, we observe annual value added and some relevant background characteristics, e.g. industry.

[21] In fact, $\xi(\mathbf{y}_0, \mathbf{y}) \equiv \frac{e(\mathbf{y}_0)p(\theta(\mathbf{y}_0))f(\mathbf{y})}{r+e(\mathbf{y}_0)p(\theta(\mathbf{y}_0))} = \left[\frac{e(\mathbf{y}_0)p(\theta(\mathbf{y}_0))}{r+e(\mathbf{y}_0)p(\theta(\mathbf{y}_0))}\right]\left[\frac{e(\mathbf{y}_0)p(\theta(\mathbf{y}_0))f(\mathbf{y})}{e(\mathbf{y}_0)p(\theta(\mathbf{y}_0))}\right]$. The second term in the right-hand side product is the density that an unemployment spell ends with the worker making a transition to a rung-$\mathbf{y}$ job. The first term is the discount factor that must be applied to a payment expected to received $1/e(\mathbf{y}_0)p(\theta(\mathbf{y}_0))$ units of time into the future, when, as is the case in our model, the duration until the payment is made follows an Exponential distribution.

6.1 Data Sources

Our data has three sources: (i) labour market spell data; (ii) IDA, a register-based matched employer-employee database maintained by Statistics Denmark; (iii) administrative firm-level VAT accounts from the Danish tax authorities. Worker and firm identifiers common to the three data sources obtains from the social security registry and the business registry, respectively.

Labour Market Spell Data. The labour 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.[22] A job spell is defined as a continuous period of primary employment at a given firm.[23] Nonemployment spells are periods where no job spells are recorded, with no distinction between unemployment and nonparticipation spells. We recode nonemployment spells shorter than 14 days between jobs at two different firms as employment at the origin firm, and also recode nonemployment spells shorter than 12 weeks between two consecutive job spells at the same firm as employment. The unit of observation in the labour market spell data is a person-spell-year. The job spell data includes worker and firm identifiers, start- and end-dates of the job, and the average annual wage rate in each job.[24]

IDA data. IDA consists of several sub-panels available from 1980 onwards. We use the sub-panels IDA-P and IDA-S. IDA-P contains annual information 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. The unit of observation in IDA-P is a person-year.

IDA-S contains information on all physical workplaces in Denmark.[25] We retain information on industry affiliation and a public sector indicator from IDA-S. Our analysis is carried out at the firm-level, and a firm may consist of several workplaces. We take a firm's industry affiliation and public sector status to be the industry and public sector status of its largest workplace. The unit of observation in the (aggregated) IDA-S panel is a firm-year.

VAT data. Firms' annual sales and purchases are obtained from the data set MOMS, constructed from firm-level VAT accounts held by the Danish tax authorities, and is available from

[22] Henning Bunzel at Aarhus University has been instrumental in developing the labour market spell data. Bunzel and Hejlesen (2016) provide a detailed description of the construction of the labour market spell data.

[23] 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. A firm may consist of multiple workplaces. Continuous employment at different workplaces within a firm is considered as a single job spell.

[24] We observe annual earnings in each job and an estimate of the annual number of hours worked in the job based on on mandatory pension contributions. Lund and Vejlin (2015) develop and implement a procedure for computing annual hours in a job in the IDA data, primarily using information on mandatory pension contributions. We adapt this procedure for the spell data with some minor simplifications.

[25] Some jobs involve work that is carried out at different and changing locations. Statistic Denmark designate such jobs as taking place at fictitious workplaces. Fictitious workplaces are excluded from IDA-S.

1990. 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 first merge IDA-P information onto the labour market spell data by person identifier and year and retain only person-years that are found in both data sources (99 percent of the spell data observations pertaining to persons aged 15-74 are matched with an IDA-P observation). Next, we merge the spell data/IDA-P intersection with the IDA-S data and the VAT data by firm identifier and year. We retain all observations in the spell data/IDA-P intersection, whether or not they are matched to an IDA-S or a VAT observation.

6.2 The Analysis Data

We restrict attention to the ten-year period 1994-2003, a period where the Danish tax system did not undergo major reforms. We discard observations on individuals never observed with either age or education information, as well as individuals with implausible education information.[26] We then define labour 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. To stay clear of behaviour driven by retirement considerations, we truncate labour market histories at the last year an individual is observed residing in Denmark or at December, 31st of the year an individual turns 55, whichever occur first. Our analysis abstracts from (extensive as well as intensive margin) labour supply responses to income taxation, with the structural model intended to represent behaviour of core labour market participants who are either in (or searching for) full time employment. We therefore discard workers who, in any year during 1994-2003, worked less than 25 hours a week on average. Finally, we delete all observations on workers who, in any year during 1994-2003, reside outside Denmark. Wage and value added are trended to 2003 prices using the internal deflator computed from repeated annual cross sections (November 28) in the analysis panel.

The selected analysis data contains 3,852,637 job spells and 1,223,989 nonemployment spells on 1,559,599 individuals involving 191,726 firms during 1994-2003. On average, an individual is present for 7.5 years and employed in 2.2 different firms during the 10-year observation period. Looking at pooled annual cross sections, the average age in the analysis data is 40 years, 40 percent are women, the employment rate is 89.3, and the public sector employment rate (as a percentage of the total labour force) is 29 percent. A later section provides detailed descriptive analysis of the analysis data in the form of the statistics used in our calibration.

6.3 The Effective Marginal Tax Rate

The actual Danish 1994-2003 tax system is progressive with an increasing marginal tax rate, and is immensely more complicated than the affine tax functions we study in this paper. Our

[26]The primary cause of missing education data is foreign educational credentials. We map information on educational attainment into years of schooling and consider observations where age minus years of schooling is less than 5 as implausible.

empirical analysis therefore relies on an affine approximation to the actual tax function. We base our approximation on a simulator of the Danish tax system developed by Kleven and Schultz (2014), which relies on rich individual-level information on types of taxable incomes, region of residence, marital status, and, if married, spousal income, information that is available from the IDA data.[27] Our approximate tax function also accounts for a Danish VAT rate of 0.25.

Given our focus on labour 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 is the actual tax function in year t that we want to approximate, LI_{it} is individual labour income, and $\mathbf{Z}_{it}$ is a vector of relevant characteristics and other income concepts relevant for the tax liabilities of the individual i in year t (the main components of which are capital income, itemized deductions, and other personal income such as benefits). We first simulate the total tax liability for each individual, for each year in the 1994-2003 observation window, and compute individual marginal tax rates as

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

where the marginal change of 100 Danish Kroner is approximately 15 US Dollars in 2003 prices.

Using the individual marginal tax rates we compute an average marginal tax rate of 0.554 across ten annual (November 28th) cross sections in our analysis data. Accounting for a 25 percent VAT rate we obtain the following estimate of the effective constant marginal tax rate t:

$$t = \frac{0.554 + 0.25}{1 + 0.25} = 0.643.$$

In an international context, the estimated tax rate is high, but it is consistent with other attempts to measure the effective marginal income tax in Denmark, see e.g. Kleven and Schultz (2014), and Kreiner, Munch, and Whitta-Jacobsen (2015).

7 Model Calibration

7.1 Parameterisation

Functional forms and distributional assumptions. We conduct the empirical analysis with $I = 10$ worker-types, indexed by i and differentiated by a "skill-level" a; hence, a_i is the skill-level of worker type-i. Let $\{a_i; i = 1, 2, \ldots, 10\}$ be ten equidistant points on the interior of the unit interval. Now, let g_a be the PDF of a Beta-distributed random variable parameterized by the two shape parameters χ_1^a and χ_2^a. The ten-point discrete (sampling) distribution of a is constructed as:

$$\Pr(a = a_i) = \frac{g_a(a_i)}{\sum_{i'=1}^{10} g_a(a_{i'})}; \quad i = 1, 2, \ldots, 10. \tag{40}$$

In our theoretical analysis we assumed job attributes to be continuously distributed to ensure

[27] We are grateful to Henrik Kleven and Esben Schultz for making their 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.

that marginal tax rate changes involved distortions to workers' ranking of jobs. Our simulation-based empirical analysis, however, requires a discrete sampling distribution of job attributes. We shall assume that each worker type faces a job attribute distribution with 10 productivity attribute-levels and 10 amenity attribute-levels, which implies 100 submarkets per worker-type. As we operate with ten worker-types, our empirical implementation thus entails segregating the labour market into a total of 1,000 submarkets.[28]

Let j index the ten productivity attribute levels and let y_p^{ij} be the productivity attribute in a match between a type-i worker and a job with level-j productivity. We assume

$$y_p^{ij} = \varrho_0 + \varrho_1 a_i + \varrho_2 p_j + \varrho_3 a_i p_j, \tag{41}$$

where p is a ten-point discrete random variable constructed by replacing g_a in (40) by the Beta PDF g_p with shape parameters χ_1^p and χ_2^p. ϱ_0 is an intercept and ϱ_1, ϱ_2, and ϱ_3 are loading parameters. Note that the sampling distribution of p is common to all worker-types. However, by appropriately restricting the loading coefficients in the match production function (41), we ensure that, effectively, high-type workers sample "better" productivity-attributes.

We further assume that different worker-types sample amenity attributes from a common distribution, and that sampled productivity and amenity attributes are independent. Let k index the assumed ten amenity attribute levels and let y_z^k be the amenity attribute in a job with level-k amenity. We then impose

$$y_z^k = \bar{z} z_k, \tag{42}$$

where $\bar{z}$ is a loading parameter and z_k is a ten-point discrete random variable constructed by replacing g_a in (40) by the Beta PDF g_z with shape parameters χ_1^z and χ_2^z.

Recall that the matching of workers and vacancies in each of the submarkets is governed by a Cobb-Douglas matching function, $m(E, V) = AE^{\beta}V^{1-\beta}$, where E and V are aggregate worker search effort and number of vacancies in the submarket, respectively.

Following Christensen, Lentz, Mortensen, Neumann, and Werwatz (2005) we assume that the search cost function $c(e)$ is a power function given by

$$c(e) = \frac{c_0}{1 + 1/c_1} e^{1+1/c_1}, \tag{43}$$

where c_0 and c_1 are parameters.

Finally, we assume that wages are observed with measurement error, which is normally distributed with mean zero and variance σ_ϵ^2.

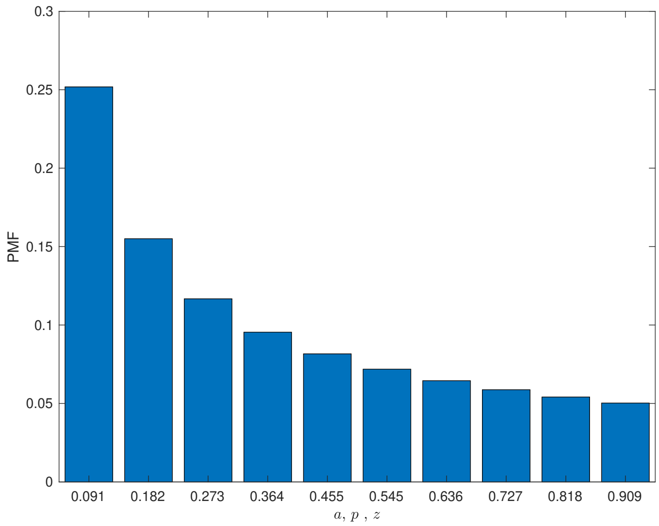
Fixed parameters. Table 1 lists the parameter values that we fix prior to the calibration. The effective annual discount rate, r, is set to 0.05. The elasticity of the matching function, β, is set to 0.5, a common choice in empirical work, see e.g. Petrongolo and Pissarides (2001).

[28] In the empirical analysis, we compute the deadweight losses by applying the formulas obtained for continuous sampling distributions, thus treating the discrete sampling distributions with 100 submarkets per worker-type as an approximation to a true continuous distributions.

The scale of the search cost function c_0 is not separately identified from matching efficiency A. We leave A free and normalize c_0 to unity. Based on Christensen, Lentz, Mortensen, Neumann, and Werwatz (2005), we set $c_1 = 1$ such that the search cost function is quadratic.

To facilitate comparison between the estimated loading coefficients in the productivity attribute and amenity attribute equations, see (41) and (42), we fix the worker type distribution and the sampling distributions of the productivity and amenity attributes to be identical; specifically, we set $\chi_1^a = \chi_1^p = \chi_1^z = 0.3$ and $\chi_2^a = \chi_2^p = \chi_1^z = 1$. These parameter values yield Beta densities that are everywhere declining. Figure 4 plots the implied distributions of a, p, and z.

Figure 4: Common distributions of a, p, and z



We set unemployment (flow) benefits, b, to 100 (Danish Kroner), which is consistent with the actual (hourly) benefit level in Denmark over our data period, and set $\varrho_0 = b$, which together with the maintained assumption that $y_z = 0$, ensures that unemployed workers accept all job offers. We furthermore impose the restriction $\varrho_3 = 0$ in (41).

The cost of entry, K, is set to 1 and we fix the flow cost of operating a vacancy, c^v, to 0.05. Finally, we set the profit tax rate to 0.2, which is consistent the actual Danish profit tax rate during our sample period and maintain the assumption that entry and vacancy operating costs are fully deductible to ensure that profit taxation is non-distortionary, see Lemma 1.

Data driven parameters. There are seven free parameters which we calibrate to match seven data moments. While the parameters are calibrated jointly, each moment is included to identify one of the free parameters.

The scale of the matching function, A, targets the probability of finding a job within 6 months of becoming unemployed.[29] The job destruction rate, s, targets the annual job destruction rate.

The calibration of the loading parameters ϱ_1 and ϱ_2 in the match production function (41), and the wage measurement error variance, σ_ϵ^2 is based on a statistical decomposition of the vari-

[29]We could alternatively target the nonemployment rate, but prefer to base our calibration on worker-flows to avoid contaminating the calibrated A by the presence of nonparticipants among the stock of nonemployed individuals in our data. Indeed, as we only impose very mild sample selection restrictions, the empirical nonemployment rate fairly high, likely for reasons unrelated to search frictions.

Table 1: Fixed Parameters

Parameter	Description	Value
r	Effective annual discount rate	0.05
β	Matching function elasticity	0.5
K	Entry cost	1
c^v	Vacancy operating cost	0.05
c_0	Scale parameter in search cost function	1
c_1	Elasticity of search cost function	1
χ_1^a	Worker-type distribution	0.3
χ_2^a	Worker-type distribution	1
χ_1^p	Productivity-attribute distribution	0.3
χ_2^p	Productivity-attribute distribution	1
χ_1^z	Amenity-attribute distribution	0.3
χ_2^z	Amenity-attribute distribution	1
b	Unemployment benefits	100
ϱ_0	Production function parameter (intercept)	100
ϱ_3	Production function parameter (complementarity)	0
τ	Profit tax rate	0.2
γ_K	Fraction deductible of K	1
γ_c	Fraction deductible of c^v	1

ance of individual log wages into a between-worker component, a within-worker, between-job component, and a within-job component.[30]. The between-worker component captures heterogeneity in worker-types, and aids identification of ϱ_1. The within-worker, between-job component captures heterogeneity in productivity-attributes for wages and aids identification of ϱ_2. Finally, within our model, wages vary within-job only due to i.i.d. measurement errors, and the within-job wage variance component thus pins down σ_ϵ.

Absent amenities, or when amenities are constant, job-to-job transitions with observed wage cuts are driven entirely by i.i.d. measurement errors in wages, which must also account for within-job wage variation. Dispersion in amenity attributes, however, provide a second source of job-to-job transitions associated with wage cuts. Thus, we fit the proportion of job-to-job transitions involving a wage cut to identify the scale of the amenity attributes, $\overline{z}$.

Finally, the tax rate is set to the average marginal tax rate in the data.

[30] Formally, the calibration of ϱ_1, ϱ_2, and σ_ϵ^2 is based on the following log wage variance decomposition:

$$\underbrace{\frac{\sum_{i=1}^{N}\sum_{j=1}^{J_i}\sum_{t=1}^{T_{ij}}(w_{ijt}-\overline{w})^2}{\sum_{i=1}^{N}\sum_{j=1}^{J_i}T_{ij}}}_{\text{Total variance}} = \underbrace{\frac{\sum_{i=1}^{N}\sum_{j=1}^{J_i}\sum_{t=1}^{T_{ij}}(w_{ijt}-\overline{w}_{ij})^2}{\sum_{i=1}^{N}\sum_{j=1}^{J_i}T_{ij}}}_{\text{Within-job}} + \underbrace{\frac{\sum_{i=1}^{N}\sum_{j=1}^{J_i}\sum_{t=1}^{T_{ij}}(\overline{w}_{ij}-\overline{w}_{i})^2}{\sum_{i=1}^{N}\sum_{j=1}^{J_i}T_{ij}}}_{\text{Within-worker, between-job}} + \underbrace{\frac{\sum_{i=1}^{N}\sum_{j=1}^{J_i}\sum_{t=1}^{T_{ij}}(\overline{w}_{i}-\overline{w})^2}{\sum_{i=1}^{N}\sum_{j=1}^{J_i}T_{ij}}}_{\text{Between-worker}},$$

where w is the log wage, $\overline{w}$ indicate average log wage, N is the number of individuals in the data (indexed by i), J_i is the number of jobs that individual i holds (indexed by j), and T_{ij} is the number of wage observations in job j for individual i (indexed by t).

7.2 Calibrated Parameter Values and Fit

Table 2 lists the free parameters, the calibrated parameter values, and the calibrated and empirical values of the seven fitted moments.

The calibrated model provides a near-perfect fit to the data moments (but we only calibrate a small number parameters in an exactly identified calibration procedure). Moreover, the calibrated parameter values are reasonable. As the underlying distributions for worker-types a, job productivity-types p, and amenity-types z are identical (and illustrated in Figure 4), the loading coefficients $\overline{z}$, ϱ_1, and ϱ_2 are measured on the same scale, and are easily comparable. Hence, $\varrho_1 = 148$ implies that the difference in the average sampled productivity-attribute between the highest worker-type, with $a = 0.909$, and the lowest worker-type, with $a = 0.091$, is 121 Danish Kroner. Similarly, $\varrho_2 = 143$ implies a difference of 117 Kroner between the most productive job-type, with $p = 0.909$, and the least productive job-type, with $p = 0.091$. To interpret $\overline{z}$, recall that unemployment is associated with zero amenities, i.e. $y_{z,0} = 0$; hence, $\overline{z} = 100$ implies that there is a 9 Kroner difference in amenities between a job with the lowest amenity-level and unemployment, while that difference is 91 Kroner between a job with the highest amenity-level and unemployment, which implies a difference of 82 Kroner between the lowest and the highest amenity levels.

Table 2: Data Driven Parameters

Par.	Description	Value	Data Moment	Model	Data
A	Matching efficiency	0.054	Pr(unemployment spell $<$ 6 mth.)	0.537	0.537
s	Job destruction rate	0.092	Average job destruction rate	0.092	0.092
ϱ_1	Worker-type loading	148	B/w-worker log wage var.	0.106	0.106
ϱ_2	Prod.-type loading	143	W/n-wrk, b/w-job log wage var.	0.020	0.020
σ_ϵ	Measurement error var.	0.100	W/n-job log wage var.	0.010	0.010
$\overline{z}$	Amenity-type loading	100	Prop. J2J transitions w/ wage cut	0.391	0.391
t	Labour income tax rate	0.643	Average marginal tax rate	0.643	0.643

Appendix F show features of the calibrated equilibrium. As expected, high-type workers exert more search effort in unemployment vis-a-vis low-type workers, and also enjoy lower tightnesses, which results in lower unemployment rates. Workers exert ten times more search effort in the worst job compared to median job, and the equilibrium employment distribution has a lot of mass on high-amenity/low-productive jobs.

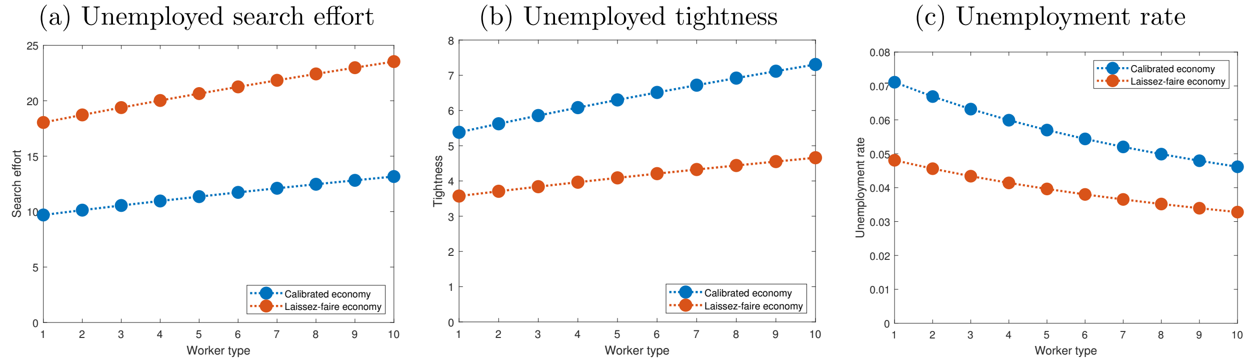
8 Quantitative Analysis

We use the calibrated model for three pieces of analysis of the distortionary effects of redistributive income taxation. We first compare the calibrated model economy to a "laissez-faire" economy without a tax and transfer system. Next, we compute and decompose the (marginal) deadweight losses associated with the tax and transfer system in the calibrated economy. Finally, we compute the optimal income tax rates implemented by a social planner with varying degrees of redistributive preferences, and analyse the deadweight losses in these optimal tax economies.

8.1 Comparison to a Laissez-faire Economy

The laissez-faire economy is without a tax and transfer system ($b = t = 0$), but is otherwise identical to the calibrated model economy.

Figure 5: Calibrated versus Laissez-faire Economy: Unemployment



Notes: The figure shows unemployed search effort and tightness in the calibrated model and in the laissez-faire model, which is the calibrated model without taxes and transfer.

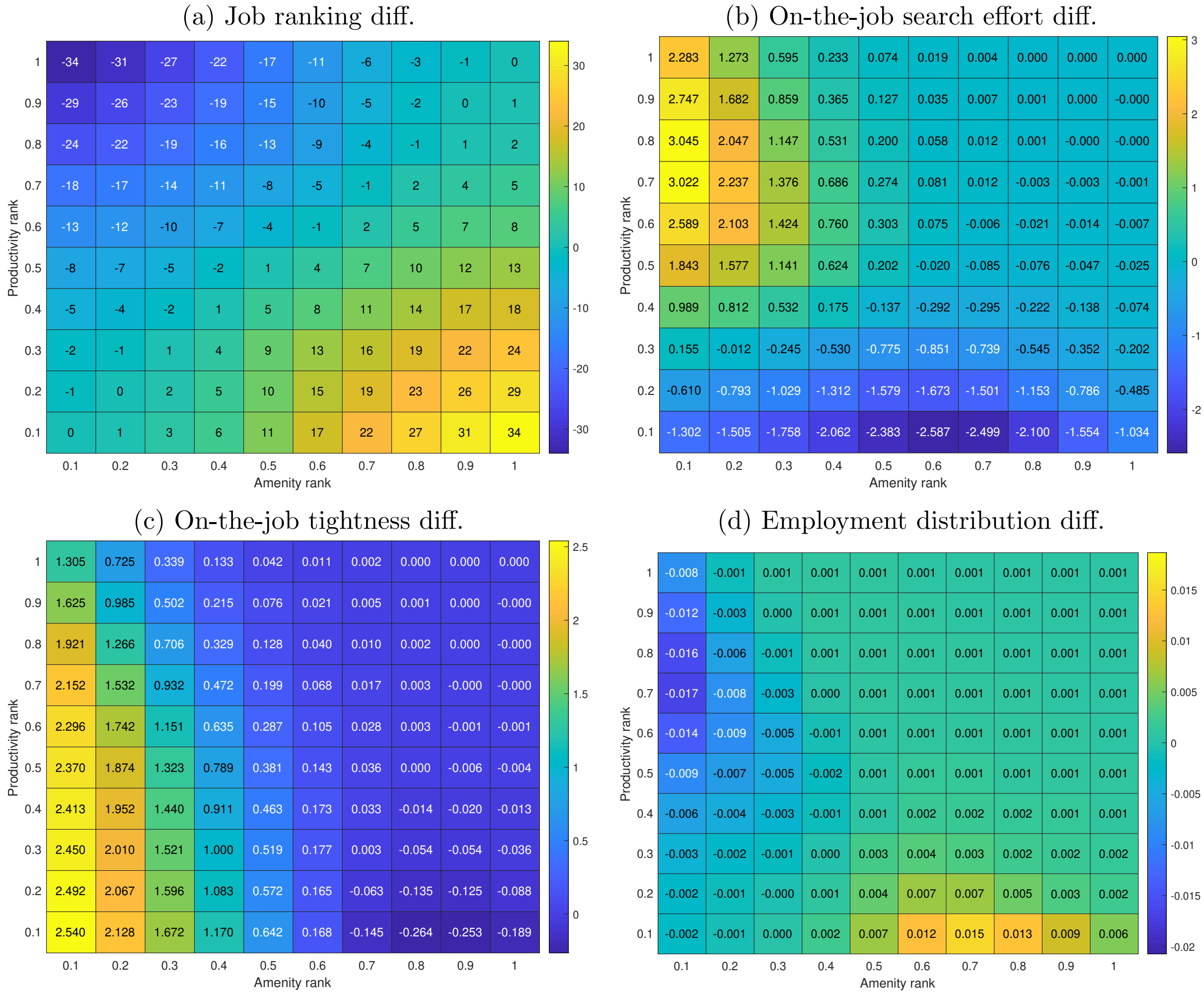
Figure 5 plots, by worker-type, unemployed search effort, tightness in the unemployment-submarkets, and the unemployment rate for the calibrated model and the laissez-faire economy. Panel (a) shows that the tax and transfer system stifles unemployment search effort (for all worker-types). Indeed, both income taxation and unemployment benefit provision reduces unemployed workers' incentives to search for a job.

Panel (b) illustrates the effect of the tax and transfer systemt on vacancy creation. Unemployment benefit provision reduces firms' incentive to post vacancies to the unemployment submarkets, which tend to lower labor market tightness in the calibrated economy vis-a-vis the laissez-faire economy. However, taxation in the calibrated economy incentivizes vacancy creation since future jobs hold promises of untaxed amenities. The lower tightness in the laissez-faire economy evident in panel (b) shows that, in the calibrated model economy, the upward pressure on the tightness coming from taxation in fact dominates the downward pressure coming from unemployment benefit provision.

Panel (c) plots the worker-type specific unemployment rates in the calibrated and in the laissez-faire economy, and shows that the effect of the lower search effort in the calibrated economy dominates the effect of inflated vacancy creation, such that the calibrated economy features higher unemployment rates than the laissez-faire economy for all worker-types.

Figure 6 pertains to the submarkets of employed workers. Panels (a), (b) and (c) show heatmaps that illustrate the difference in employed workers' ranking of jobs, search effort, and the tightnesses they face between the calibrated model economy and the laissez-faire economy. Positive (negative) entries implies that a particular submarket has a higher (lower) ranking, search effort, or labour market tightness in the calibrated model economy than in the laissez-faire economy. In the heatmaps, larger positive entries are indicated by warmer colours while

Figure 6: Calibrated versus Laissez-faire Economy: Employment



Notes: The heatmaps in panels (a), (b), (c), and (d) shows differences between the calibrated model economy than in the laissez-faire economy. The laissez-faire economy is the calibrated model without taxes and transfer. Positive (negative) entries implies that a particular submarket has a higher (lower) value in the calibrated model economy than in the laissez-faire economy.

larger negative entries are indicated by colder colours.[31]

Panel (a) in Figure 6 shows that workers in the laissez-faire economy tend to rank high-productive, low-amenity (low-productive, high-amenity) jobs higher (lower) than workers in the calibrated model economy.[32] The same line of logic applies to panel (b), which shows that workers in high-productivity, low-amenity jobs (the north-west corner) exert more search effort in the calibrated model economy than in the laissez-faire counterfactual. The opposite holds true for workers in low-productivity, high-amenity jobs (the south-east corner).

Panel (c) in Figure 6 shows the difference in tightness faced by employed workers in the estimated and the laissez-faire economy. Most of the entries in panel (c) are positive, and they are particularly large in submarkets with low amenities. This suggest that the tax and transfer

[31]The laissez-faire economy retains the property from the calibrated model economy that the behaviour of employed workers is independent of worker-type; hence, Figure 6 is unconditional on worker-type.

[32]Recall that there are 100 markets for each worker type, so the ranking takes values from 1 to 100, with the rank-1 submarket being the least attractive and the rank-100 submarket being the most attractive. In the heatmap in Figure 6, panel (a), an entry of, say, 27, indicates that workers in the calibrated model economy rank the specific submarket 27 positions higher than workers in the laissez-faire economy.

system boosts vacancy creation, and particularly so in low-amenity submarkets. Indeed, workers in low amenity jobs expect higher future amenity values ($\Delta Y_z(y) > 0$), and the tax system in the calibrated model economy incentivises vacancy creation in these submarkets. In submarkets where the productivity-attribute is low relative to the amenity-attribute (i.e. in the south-east corner), workers expect to ascend the productivity-attribute ladder and descend the amenity-attribute ladder, and the tax system disincentivises vacancy creation.

Panel (d) in Figure 6 shows how differences in job ranking, search effort, and vacancy creation between the calibrated model and the laissez-faire economy result in different employment distributions. It is evident that the tax and transfer system reallocates employment away from high-productive, low-amenity jobs towards low-productive, high-amenity jobs.

8.2 The Deadweight Losses in the Calibrated Economy

We compute and decompose the marginal deadweight loss in the calibrated economy according to (38) and (39). The results are presented in Table 3, where column (1) reports marginal deadweight loss components as a proportion of the total marginal deadweight loss, and column (2) reports marginal deadweight loss components as a proportion of the tax base $\overline{W}_0$.

Table 3 shows that the marginal deadweight loss represents 33 percent of the tax base, and that distortions to job search effort, job ranking, and vacancy creation comprises 28 percent, 7 percent, and 65 percent of the marginal deadweight loss, respectively.

The total deadweight loss is also split into components coming from the revenue- and the expenditure-side of the government budget. Table 3 shows that the revenue-side contributes 111 percent of the total deadweight loss through distortions to job search effort, job ranking, and vacancy creation on the job ladder, all of which shrinks the tax base. Consequently, the expenditure-side, i.e. unemployment benefit provision, contributes negatively to the total marginal deadweight loss, namely -11 percent, which we decompose into components stemming from distorted search and vacancy creation in the unemployment submarkets. With a tax and transfer system, unemployed search effort is lower than the planner's stipulation. A higher tax rate reduces unemployed search even further, which increases aggregate unemployment expenditures and the total marginal deadweight loss (by 7 percent). Unemployment benefit provision decreases vacancy creation in the unemployment submarket below the planner's stipulation, but higher taxation (in the presence of amenities and provided unemployment is at the bottom of the amentity rung) in fact increases vacancy creation and restores some of the efficiency loss from unemployment benefits provision (namely -18 percent).

The elasticity of the tax base $\overline{W}_0$ with respect to the tax rate t is a key component in the optimal tax design, see (26). The calibrated economy has $t = 0.643$, which implies an elasticity of the tax base with respect to the income tax rate of $0.365 \times 0.643 = 0.234$, where 0.365 is the revenue-side deadweight loss as a proportion of the tax base $\overline{W}_0$, see Table 3, column (2).

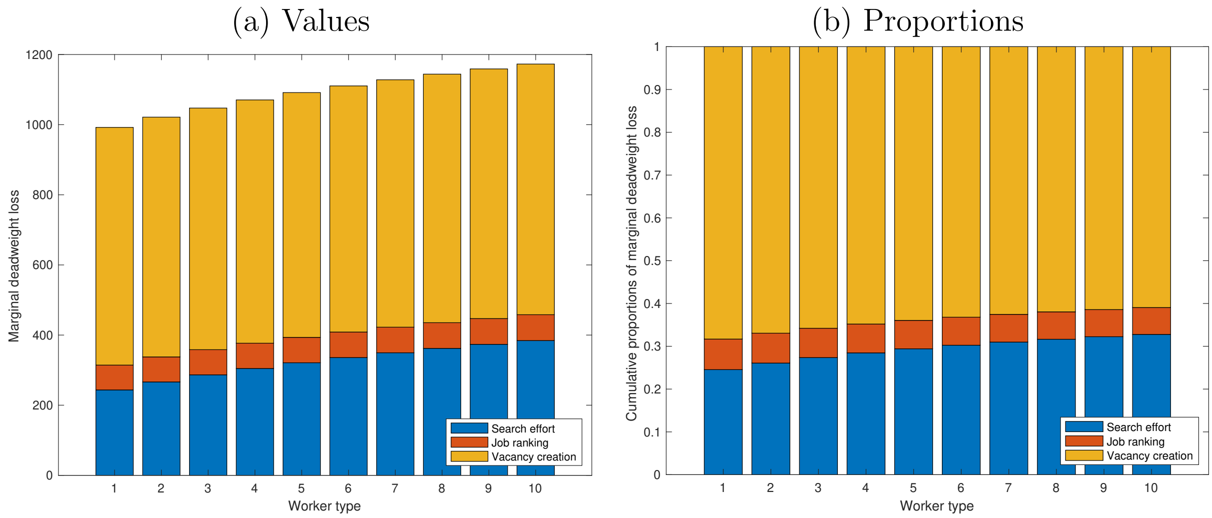
The model framework allow us to further expose the structure and sources of the deadweight losses. First, Figure 7 reports the marginal deadweight losses and their composition by worker-type. Panel (a) shows that the deadweight loss is increasing in worker-type, but not dramatically

Table 3: The Marginal Deadweight Losses from Taxation in the Calibrated Economy

	(1)	(2)
	Proportion of marginal deadweight loss	Proportion of tax base, $\overline{W}_0$
Total	1.000	0.328
Job search effort distortions	0.281	0.092
Job ranking distortions	0.068	0.022
Vacancy creation distortions	0.651	0.214
Revenue side, total	1.111	0.365
Job search effort distortions	0.210	0.069
Job ranking distortions	0.068	0.022
Vacancy creation distortions	0.833	0.274
Expenditure side, total	−0.111	−0.037
Job search effort distortions	0.071	0.023
Vacancy creation distortions	−0.182	−0.060

so. Panel (b), which plots the proportional contribution of job search effort, job ranking and vacancy creation for each worker-type, shows that distortions to job search effort is relatively more important for high-type workers, while distortions to vacancy creation is relatively more important for low-type workers; however, the difference are rather small.

Figure 7: The Marginal Deadweight Loss by Worker-Type



In Appendix F we show job ladder rung-specific single-spell partial deadweight losses $O_t^e(\mathbf{y})$, $O_t^\phi(\mathbf{y})$ and $O_t^\theta(\mathbf{y})$ from distorted job search, job ranking, and vacancy creation in unemployment, all 100 employment states, and (where relevant) for different worker-types, cf. Section 5. Roughly speaking, the marginal deadweight loss contributions are large in the submarkets where the difference in the variable in quesion between the calibrated and the Laissez-faire economy is large, see Figure 6, panels (a)-(c).[33]

8.3 Optimal Redistributive Taxation

The planner levies taxes on labour income to maximize aggregate welfare defined by (20). We parameterise the planner's redistributive preferences by the CRRA function,

$$\Phi(x) = \frac{x^{1-\gamma} - 1}{1 - \gamma}, \tag{44}$$

and refer to γ as the planner's degree of inequality aversion. We compute the optimal tax t associated with different values of γ. In these computations, we hold the level of benefits constant at the calibrated level, $b = 100$ and take $\overline{G} = 0$.[34]

Panel (a) in Figure 8 plots (in green) the optimal income tax rate, t, as a function of the planner's inequality aversion, γ. The tax rate in the calibrated economy, $t = 0.643$, is also indicated. The optimal income tax rate t is naturally increasing in γ and an even moderately inequality averse planner taxes labour income at nontrivial rates; for example, for $\gamma = 1$, where $\Phi(x) = \log(x)$, the optimal income tax rate is around 25 percent.[35] The optimal income tax rate, however, does not come close to the actual income tax rate of 64 percent for any reasonable value of γ. Allowing for exogenous government spending, i.e. taking $\overline{G} > 0$, will yield a closer alignment of optimal and actual tax rates.

Panel (b) in Figure 8 plots the coefficient of variation of V_0^i across worker-types as a function of γ, and also indicate the coefficient or variation of V_0^i for the calibrated economy. Of course, V_0-dispersion declines with γ as a result of the redistribution implemented by the increasingly inequality averse planner (the average V_0 remains effectively constant). The calibrated economy has a coefficient of variation of 0.062, which is substantially lower than 0.096, the coefficient of variation in the economy designed by the very inequality-averse planner with $\gamma = 5$.

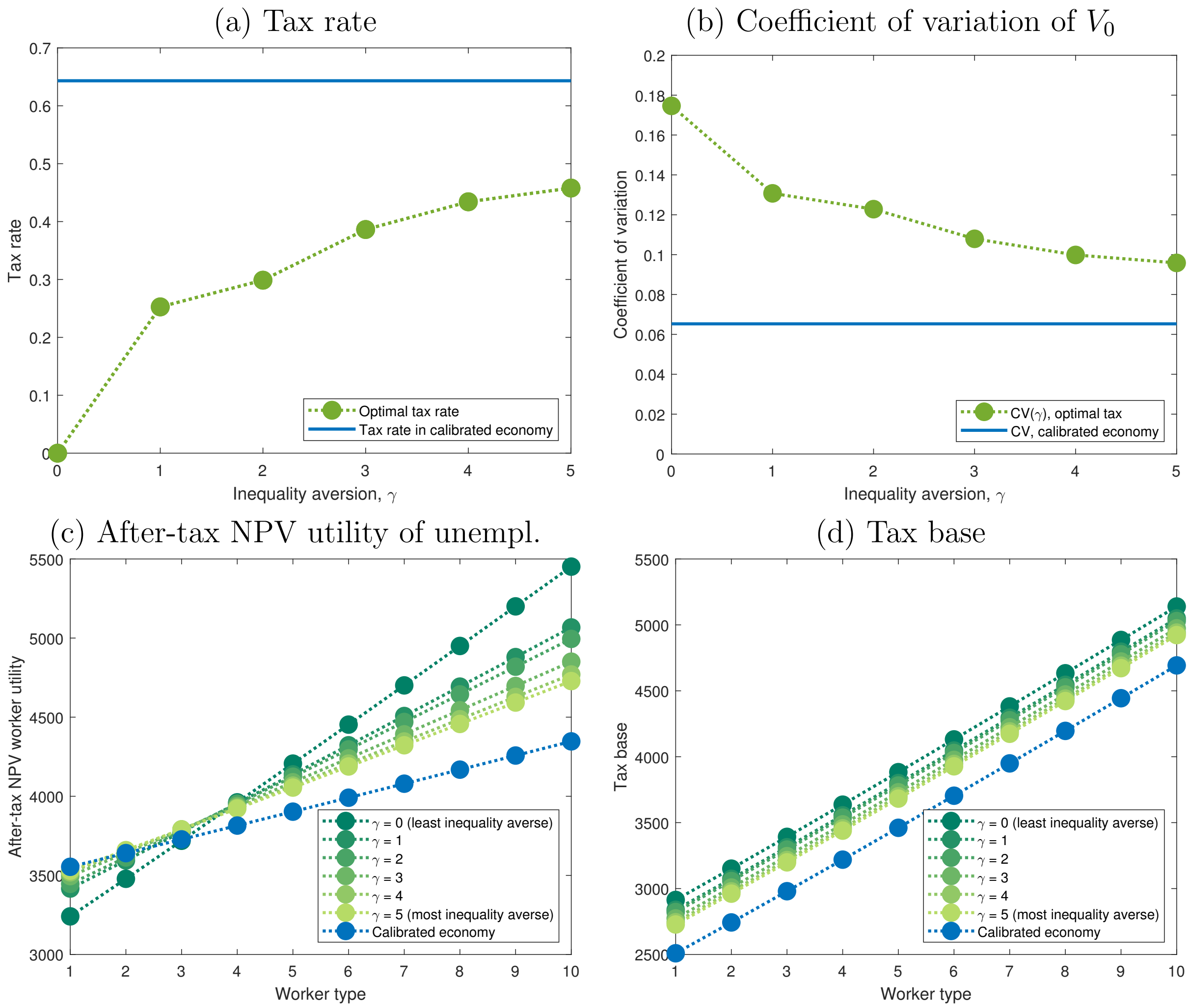
Panels (c) and (d) in Figure 8 plot the after-tax NPV utility V_0^i and tax base W_0^i as a function of worker-type for different values of γ, and for the calibrated economy. In panel (c), Moving from the inequality-indifferent planner economy ($\gamma = 0$) to economies with redistributive policies in place ($\gamma > 0$), benefits lower-type workers at the expense of higher-type workers. Indeed, depending the value of γ, the two or three least productive worker-types emerge as net-benefactors. Panel (d) confirms that the tax base decreases with γ, i.e. that the deadweight loss increases with γ. Evidently, the deadweight loss in the calibrated economy is substantially higher than the deadweight loss incurred even by an extremely inequality averse planner, consistent with lower levels of redistribution in the planner economy than in the calibrated economy. We return to this issue further below.

[33] Appendix F also shows the weights ω_0^i and $\omega_1^i(\mathbf{y})$ that are used in the aggregation of the single-spell partial deadweight losses, see (39).

[34] Note that $\overline{G} = 0$ is equivalent to assuming that government exogenous spending is evenly distributed across workers and a perfect substitute to private income. In that case, exogenous government spending can be subsumed in the lump sum transfer t_0.

[35] An inequality-indifferent planner with $\gamma = 0$ set $t = 0$, as in the laissez-faire economy, and finance the (fixed) unemployment benefits by lump sum taxation.

Figure 8: Taxation in Planner Economies and the Calibrated Economy



Notes: The parameter γ index the planner's "inequality aversion" in the CRRA utility weight, (44). The average V_0 remains effectively constant for different values of γ.

8.3.1 Deadweight Losses in the Optimal Tax Economies

The tax rate in the calibrated economy is $t = 0.643$, a substantially heaver tax burden than implemented by a very inequality averse planner. Figure 9 plots the marginal deadweight losses, and its decomposition into components stemming from distorted search effort, job ranking and vacancy creation, for the calibrated economy and for the optimal tax economies discussed above.

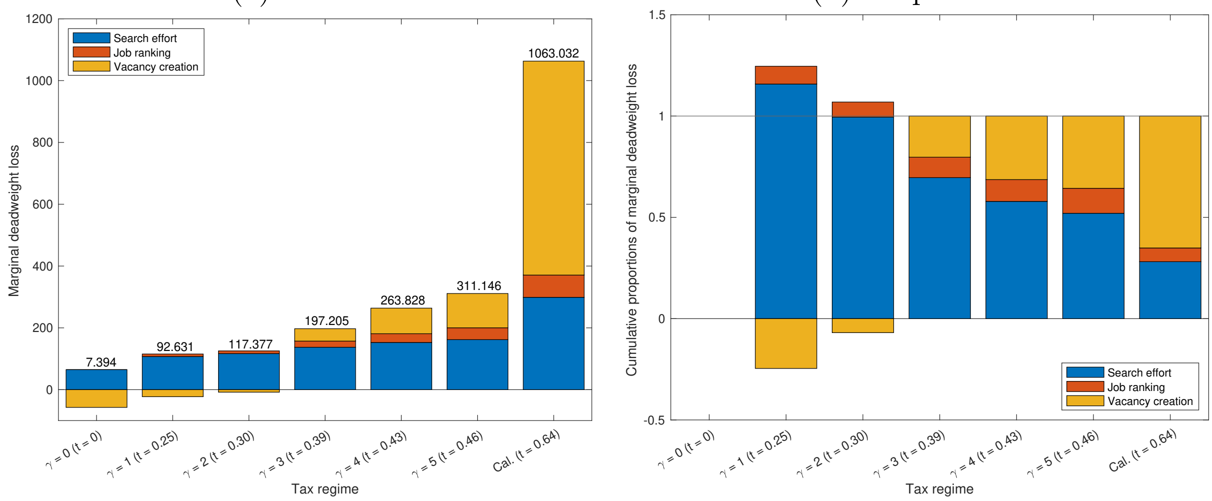
Panel (a) in Figure 9 shows the marginal deadweight losses for income tax rates optimally set by planners with varying degrees of inequality aversion as indexed by the parameter γ in (44). RWe hold the level of unemployment benefits fixed in the calibrated and optimal tax economies. The inequality-indifferent planner with $\gamma = 0$ sets $t = 0$ and finance unemployment transfers by lump sum taxes. The provision of unemployment benefits generates deadweight losses in the model. As a consequence the marginal deadweight loss from increasing taxation from a level of zero has first order effects. In particular, it gives rise to a small deadweight loss of 7 Kroner, which is the net effect of a positive deadweight loss coming from distorted job search effort, and a negative deadweight loss coming from excessive vacancy creation for reasons discussed above.

For planners with higher degree of inequality aversion, the optimal tax rate increases, as does the marginal deadweight loss. All constituent components of the marginal deadweight loss increases with γ (and t); in particular, the deadweight loss coming from distorted vacancy creation turns positive once we consider the $\gamma = 3$ planner who implements $t = 0.39$. We see, however, that the deadweight losses in the calibrated economy with $t = 0.64$, which is 1,063 Kroner, is substantially higher than the deadweight losses even a very inequality averse planner with $\gamma = 5$ is willing to incur, namely 311 Kroner. Panel (b) in Figure 9 shows that the rising deadweight loss from additional redistribution in the job ladder economy is associated with an increase in the relative importance of deadweight losses stemming from distorted vacancy creation.

Overall, the plots in Figure 9 suggests that it is the (steeply) rising cost of distorted vacancy creation that makes even very inequality-averse planners abstain from implementing the high tax rates and the low level of after-tax inequality from the calibrated economy.

Figure 9: The Marginal Deadweight Losses from Taxation by Tax Regime

(a) Values (b) Proportions



Notes: The numbers on top of the bars in panel (a) report the total marginal deadweight losses. We do not report the proportions of search effort and vacancy creation distortions in the total marginal deadweight loss for the $t = 0$ tax regime in panel (b). The marginal deadweight loss in this regime is small, and as it is the net effect of a positive deadweight loss from search effort and a negative deadweight loss from vacancy creation of almost equal magnitude, proportions fall way outside the unit interval.

9 Conclusion

We have analysed as yet overlooked deadweight losses of taxation in labor markets with frictions arising through labor misallocation. To this end, we constructed a rich job ladder model of competitive on-the-job search, which includes heterogeneous workers and heterogeneous jobs. A job comprises both a productivity component and a non-taxed amenity component. The competitive search equilibrium allocation maximizes the workers' after-tax NPV utilities given the resource constraints in the economy and the policy environment; nonetheless, fiscal externalities associated with income taxation and unemployment benefit provision gives rise to deadweight losses. The deadweight losses arise from distortions to three margins: workers' job search effort, their ranking of jobs, and firms' vacancy creation. We derive analytical expressions for the marginal deadweight losses from each of these distortions at each rung of the job ladder.

We calibrate the model using Danish matched employer-employee data, and find that the marginal deadweight losses from taxation evaluated at the actual average marginal tax rate of 0.643 comprises 33 percent of the average NPV income of unemployed workers (the tax base); 28 percent of the marginal deadweight loss can be attributed to distorted search effort, 7 percent to distorted job ranking, and 65 percent to distorted vacancy creation.

Finally, we derive the optimal tax rate, i.e. the tax rate set by benevolent planner, as a function of that planner's inequality aversion. We find that the actual tax rate is substantially higher than the optimal tax rate, even if the planner has a very high level of inequality aversion. Our analysis suggest that the planners' redistrbution policies are constrained by steeply rising deadweight losses from distorted vacancy creation.

Our paper contains some loose ends that we would like to address in the future. First, our calibration can be improved, particularly when it comes to the calibration of the production function and of government expenditures. Second, it would be interesting to extend our model to allow for non-linear taxes on labour income, profit taxation, and to include *ex ante* heterogeneous firms. Finally, we would like to derive the optimal level of unemployment benefits as a function of the tax rate and the planner's inequality aversion.

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.

Best, M., and H. Kleven (2013): "Optimal Income Taxation with Career Effects of Work Effort," mimeo.

Boone, J., and L. Bovenberg (2002): "Optimal labour taxation and search," *Journal of Public Economics*, 85(1), 53–97.

Breda, T., L. Haywood, and H. Wang (2019): "Labor Market Responses to Taxes and Minimum Wage: A Dynamic Structural Approach," mimeo.

Bunzel, H., and M. Hejlesen (2016): "Documentation Spell Dataset," mimeo.

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

Cahuc, P., F. Postel-Vinay, and J.-M. Robin (2006): "Wage Bargaining with On-the-Job Search: Theory and Evidence," *Econometrica*, 74(2), 323–364.

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).

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

Garibaldi, P., E. R. Moen, and D. E. Sommervoll (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.

Guerrieri, V., R. Shimer, and R. Wright (2010): "Adverse Selection in Competitive Search Equilibrium," *Econometrica*, 78(6), 1823–1862.

Hall, R. E., and A. I. Mueller (2018): "Wage Dispersion and Search Behavior: The Importance of Nonwage Job Values," *Journal of Political Economy*, 126(4), 1594–1637.

Hosios, A. J. (1990): “On the Efficiency of Matching and Related Models of Search and Unemployment,” *Review of Economic Studies*, 57, 279–98.

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.

Hwang, H.-s., D. T. Mortensen, and W. R. Reed (1998): “Hedonic Wages and Labor Market Search,” *Journal of Labor Economics*, 16(4), 815–847.

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.

Kreiner, C. T., J. R. Munch, and H. J. Whitta-Jacobsen (2015): “Taxation and the long run allocation of labor: Theory and Danish evidence,” *Journal of Public Economics*, 127, 74–86.

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

Mancino, A., and J. Mullins (2020): “Frictional Adjustment to Income Tax Incentives: An Application to the Earned Income Tax Credit,” mimeo.

Menzio, G., and S. Shi (2010): “Block Recursive Equilibria for Stochastic Models of Search on the Job,” *Journal of Economic Theory*, 145(4), 1453–1494.

Moen, E., and A. Rosen (2011): “Incentives in Competitive Search Equilibrium,” *Review of Economic Studies*, 78, 733–761.

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

Moen, E. R., and Å. Rosén (2004): “Does Poaching Distort Training?,” *The Review of Economic Studies*, 71(4), 1143–1162.

Petrongolo, B., and C. A. Pissarides (2001): “Looking into the Black Box: A Survey of the Matching Function,” *Journal of Economic Literature*, 39(2), 390–431.

Pissarides, C. A. (1985): “Taxes, Subsidies and Equilibrium Unemployment,” *The Review of Economic Studies*, 52(1), 121–133.

Pissarides, C. A. (1998): “The impact of employment tax cuts on unemployment and wages; the role of unemployment benefits and tax structure,” *European Economic Review*, 42(1), 155–183.

Pissarides, C. A. (2000): *Equilibrium Unemployment Theory*. MIT press.

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 Market Search 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.

Smith, E. (1994): "Taxation in a two-sided search model," *Journal of Public Economics*, 54(3), 423–435.

Sorkin, I. (2018): "Ranking Firms Using Revealed Preference," *The Quarterly Journal of Economics*, 133(3), 1331–1393.

Stokey, N., and R. Lucas (1996): *Recursive Methods in Economic Dynamics*. Harvard University Press, Cambridge, Mass.

Taber, C., and R. Vejlin (2020): "Estimation of a Roy/Search/Compensating Differential Model of the Labor Market," *Econometrica*, 88(3), 1031–1069.

Wilemme, G. (2021): "Optimal Taxation to Correct Job Mismatching," *Review of Economic Dynamics*, 40, 170–197.

Appendix

A Proof of Proposition 1

We have already shown that the search contract maximizes $L(\mathbf{y})$, see (14). In particular, a worker searches to maximize $L(\mathbf{y})$, and the first order condition for this maximization problem is (17). As in the main text, we represent the acceptance set $\mathscr{Y}_a(\mathbf{y}^\ell)$ by the amenity reservation function $y_z = \phi(y_p; \mathbf{y}^\ell)$. The firm solves the problem

$$\max_{\theta,\overline{V},\phi} q(\theta) \Pr\left(y_z \geq \phi(y_p;\mathbf{y}^\ell)\right) \frac{1-\tau}{1-t}\left[E^{\mathbf{y}|y_z \geq \phi(y_p;\mathbf{y}^\ell)} L(\mathbf{y}) - \overline{V}\right]$$

subject to

$$(r+s)L(\mathbf{y}^\ell) = (1-t)y_p + y_z + ep(\theta)\Pr\left(y_z \geq \phi(y_p;\mathbf{y}^\ell)\right)\left[\overline{V}(y^\ell) - L(\mathbf{y}^\ell)\right].$$

Setting up the Lagrangian and taking derivatives with respect to $\overline{V}(y^\ell)$, θ, and $\phi(y_p;\mathbf{y}^\ell)$ give the following first order conditions

$$q(\theta)\frac{1-\tau}{1-t} = \lambda e p(\theta), \tag{A1}$$

$$-q'(\theta)\frac{1-\tau}{1-t}\left(E^{\mathbf{y}|y_z \geq \phi(y_p;\mathbf{y}^\ell)} L(\mathbf{y}) - \overline{V}(y^\ell)\right) = \lambda e p'(\theta)\left(\overline{V}(y^\ell) - L(\mathbf{y}^\ell)\right), \tag{A2}$$

$$-q(\theta)\frac{1-\tau}{1-t}\left(L(y_z, y_p(y_z)) - \overline{V}(y^\ell)\right) = \lambda e p(\theta)\left(\overline{V}(y^\ell) - L(\mathbf{y}^\ell)\right). \tag{A3}$$

From (A1) it follows that $\lambda = \frac{1-\tau}{1-t}\frac{q(\theta)}{ep(\theta)}$. Substituted into (A2), and using that the elasticity of $q(\theta)$ and $p(\theta)$ are $-\beta$ and β, respectively, we obtain

$$\overline{V}(\mathbf{y}^\ell) = (1-\beta)L(\mathbf{y}^\ell) + \beta E^{\mathbf{y}|y_z \geq \phi(y_p;\mathbf{y}^\ell)} L(\mathbf{y}). \tag{A4}$$

This shows the second part of Proposition 1. Inserting for λ into (A3) gives that $L(y_z, y_p(y_z)) = L(\mathbf{y}^\ell)$. Since future productivity draws are independent of the current productivity draw, this equation implies that $y_z + (1-t)y_p(y_z) = y_z^\ell + (1-t)y_p^\ell$. This shows the first point in Proposition 1 and (15). Inserting (A4) into (14) gives that

$$E^{\mathbf{y}|y_z \geq \phi(y_p;\mathbf{y}^\ell)} J(\mathbf{y};\mathbf{y}^\ell) = (1-\beta)\frac{1-\tau}{1-t}\left(E^{\mathbf{y}|y_z \geq \phi(y_p;\mathbf{y}^\ell)} L(\mathbf{y}) - L(\mathbf{y}^\ell)\right)$$

Together with (5) and (6) this gives (18), which completes the proof.

B Bargaining

This appendix shows that the competitive search equilibrium allocation (Proposition 1) also materializes when wages and search effort are determined by bargaining under the Hosios condition. The Hosios condition stipulates that the worker's share of the match surplus is equated to β, i.e. to the absolute value of the elasticity of the vacancy filling rate with respect to labour market tightness, in every submarket, as in Hosios (1990).

If the worker receives the entire income in a type-$\mathbf{y}^\ell$ job, the NPV income to this worker is $L(\mathbf{y}^\ell)$. We assume that a worker leaves a type-$\mathbf{y}^\ell$ job for a type-$\mathbf{y}$ job if and only if $L(\mathbf{y}) > L(\mathbf{y}^\ell)$. This makes the job acceptance set $\mathscr{Y}_a(\mathbf{y}^\ell)$ the same under wage bargaining as in the competitive search equilibrium. This assumption is consistent with the bargaingin procedure in Cahuc, Postel-Vinay, and Robin (2006), in which, after successful on-the-job search, the poaching firm and the incumbent employer compete for the worker in a Bertrand game. It follows that the worker transits to the poaching firm if and only if $(1-t)y_p + y_z > (1-t)y_p^\ell + y_z^\ell$, i.e. if $L(\mathbf{y}) > L(\mathbf{y}^\ell)$.[36]

Furthermore, and again following Cahuc, Postel-Vinay, and Robin (2006), $L^i(\mathbf{y}^\ell)$ is the worker's outside option in the bargaining game. The agents bargain over gross wages, w, and search effort, e, and the solution maximizes the Nash product

$$[V^i(\mathbf{y},\mathbf{y}^\ell) - L^i(\mathbf{y}^\ell)]^\beta J^i(\mathbf{y},\mathbf{y}^\ell)^{1-\beta},$$

where the notation hides that both $V^i(\mathbf{y},\mathbf{y}^\ell)$ and $J^i(\mathbf{y},\mathbf{y}^\ell)$ depend on the wage w and the worker's search effort e. The outside option $L^i(\mathbf{y}^\ell)$ is, of course, independent of w and e in the prospective match with the type-$\mathbf{y}$ job.

The first order condition for the gross wage w reads

$$\frac{V(\mathbf{y},\mathbf{y}^\ell) - L(\mathbf{y}^\ell)}{J(\mathbf{y},\mathbf{y}^\ell)} = \frac{\beta}{1-\beta}\frac{1-t}{1-\tau}, \tag{B1}$$

which defines the wage w as a function of employer-attributes $\mathbf{y}$ and the attributes of the worker's most recent previous job (or unemployment, as might be the case), i.e. $\mathbf{y}^\ell$.

We can use the definition of $L(\mathbf{y})$ in (3) to substitute out $J(\mathbf{y},\mathbf{y}^\ell)$ in (B1), which yields

$$V(\mathbf{y},\mathbf{y}^\ell) - L(\mathbf{y}^\ell) = \beta[L(\mathbf{y}) - L(\mathbf{y}^\ell)]. \tag{B2}$$

Equation (B2) shows that the bargaining results in the worker receiving their outside option plus a share β of the match surplus, $L(\mathbf{y}) - L(\mathbf{y}^\ell)$, obtained with the type-$\mathbf{y}$ job. Importantly, the sharing rule implied by the bargaining game is consistent with (19), a sharing rule that implements competitive search allocation. Furthermore, if we substitute (B2) into (4), we obtain (16).

The first order condition for search effort e reads

$$\frac{V(\mathbf{y},\mathbf{y}^\ell) - L(\mathbf{y}^\ell)}{J(\mathbf{y},\mathbf{y}^\ell)} = -\frac{\beta}{1-\beta}\frac{\partial V(\mathbf{y},\mathbf{y}^\ell)/\partial e}{\partial J(\mathbf{y},\mathbf{y}^\ell)/\partial e}. \tag{B3}$$

Substituting the wage equation (B1) into (B3) yields

$$\frac{\partial V(\mathbf{y},\mathbf{y}^\ell)}{\partial e} + \frac{1-t}{1-\tau}\frac{\partial J(\mathbf{y},\mathbf{y}^\ell)}{\partial e} = 0. \tag{B4}$$

Using (3), it is easy to verify that (B4) coincides with the first order condition from the problem of (directly) maximizing $L(\mathbf{y})$ with respect to e. Using the recursive expression for $L(\mathbf{y})$ in (4), we can obtain workers' search effort under wage bargaining as the solution to

$$c'(e) = p(\theta)\beta \mathrm{E}^{\mathbf{y}'\in\mathscr{Y}_a}(L(\mathbf{y}') - L(\mathbf{y})),$$

which coincides with (17)

Finally, free entry of firms in each submarket ensure that the labour market tightness in each submarket is pinned down by (6) and (18).

It follows that the equilibrium allocation with bargaining under the Hosios condition is identical to the competitive search equilibrium allocation. The wage distribution under Nash bargaining is given by (19), which is also consistent with wage setting under competitive search.

C Worker Stocks and Flows

Let $n^i(\mathbf{y})$ be the measure of workers of type i in type-$\mathbf{y}$ jobs, such that $\int_{\mathscr{Y}} n^i(\mathbf{y})d\mathbf{y} = \kappa_i$.

[36] We assume that a new firm only starts competing for the worker if it is more attractive than the incumbent firm. Thus, there is no rent extraction reasons for job search as in e.g. Cahuc, Postel-Vinay, and Robin (2006).

For unemployment-submarkets, aggregate consistency imposes

$$e^i(\mathbf{y}_0)p(\theta^i(\mathbf{y}_0)) = s\int_{\mathscr{Y}\setminus\{\mathbf{y}_0\}} n^i(\mathbf{y})d\mathbf{y}, \tag{C1}$$

for every $(i, \mathbf{y}_0)$-submarket. The left-hand side of (C1) represents the measure of type-i workers leaving unemployment. They do so when they find a job that is preferable to unemployment. The right-hand side of (C1) represents the inflow, generated by job destructions, which occur at rate s

For employment-submarkets, aggregate consistency imposes

$$\left[s + e^i(\mathbf{y})p(\theta^i(\mathbf{y}))\Pr\left(\mathbf{y} \in \mathscr{Y}_a(\mathbf{y})\right)n^i(\mathbf{y})\right] = \int_{\mathscr{Y}} e^i(\mathbf{y}^\ell)p(\theta^i(\mathbf{y}^\ell))f^i(\mathbf{y})\Pr\left(\mathbf{y} \in \mathscr{Y}_a(\mathbf{y}^\ell)\right)n^i(\mathbf{y}^\ell)d\mathbf{y}^\ell, \tag{C2}$$

for every $(i, \mathbf{y})$-submarket with $\mathbf{y} \neq \mathbf{y}_0$. The left-hand side of (C2) is the measure of workers leaving the $(i, \mathbf{y})$-submarket. They do so because of job destruction, an event that occurs at rate s, or because they find a job on a higher rung of the job ladder, which happens at rate $e^i(\mathbf{y})p(\theta^i(\mathbf{y}))\Pr\left(\mathbf{y} \in \mathscr{Y}_a(\mathbf{y})\right)$. The right-hand side of (C2) represents the inflow to submarket $(i, \mathbf{y})$, made up of workers at lower rung submarkets, including unemployment, conducting successful job search. For example, the $n^i(\mathbf{y}^\ell)$ workers in the $(i, \mathbf{y}^\ell)$-submarket meet vacancies and realize a job of type $\mathbf{y}$ at rate $e^i(\mathbf{y}^\ell)p(\theta^i(\mathbf{y}^\ell))f^i(\mathbf{y})$. They reallocate to the type-$\mathbf{y}$ job if it is preferred to their current type-$\mathbf{y}^\ell$ job.

D Deriving $O_t^e(\mathbf{y})$, $O_t^\phi(\mathbf{y})$ and $O_t^\theta(\mathbf{y})$

We use the following results and remarks extensively:

Lemma 2 *The effect on $L(\mathbf{y})$ of a marginal change in t is equal to the partial effect of a marginal increase in t, keeping the endogenous variables $\theta(\mathbf{y})$, $e(\mathbf{y})$ and $\phi(y_p', \mathbf{y})$ constant.*

Proof. Since $\theta(\mathbf{y})$, $e(\mathbf{y})$ and $\phi(y_p', \mathbf{y})$ maximize $L(\mathbf{y})$, the Envelope Theorem applies. ■

The NPV of unemployment benefits to a worker at rung-$\mathbf{y}$ of the job ladder is

$$B(\mathbf{y}) \equiv \begin{cases} B_0 & \text{if } \mathbf{y} = \mathbf{y}_0, \\ B_1 & \text{if } \mathbf{y} \neq \mathbf{y}_0. \end{cases} \tag{D1}$$

where $B_0 = \frac{(r+s)b}{r[r+s+e(\mathbf{y}_0)p(\theta(\mathbf{y}_0))]}$ and $B_1 = \frac{s}{r+s}B_0$ as defined in the main text, see (9).

Remark 2 extends Remark 1 to $Y_p(\mathbf{y})$ and $B(\mathbf{y})$

Remark 2 *Let $X \in \{Y_p, B\}$. $X(\mathbf{y})$ can be represented as $X(\mathbf{y}) = \tilde{X}(e(\mathbf{y}), \phi(\cdot, \mathbf{y}), \theta(\mathbf{y}), X(\cdot))$, where $e(\mathbf{y})$, $\phi(\cdot, \mathbf{y})$ and $\theta(\mathbf{y})$ are search effort, reservation amenity function, and labour market tightness in the current rung-$\mathbf{y}$ spell, and $\tilde{X}$ is a functional. $\tilde{Y}_p$ is implicitly defined by (11), and $\tilde{B}$ is implicitly defined by $rB_0 = b + e(\mathbf{y}_0)p(\theta(\mathbf{y}_0))[B_1 - B_0]$ and $rB_1 = s[B_0 - B_1]$.*

Lemma 3 *The single-spell partial derivative of the before-tax match value $M(\mathbf{y})$ with respect to $x(\mathbf{y}) \in \{e(\mathbf{y}), \phi(y_p', \mathbf{y}), \theta(\mathbf{y})\}$ is*

$$\frac{\partial \tilde{M}(\mathbf{y})}{\partial x(\mathbf{y})} = t\frac{\partial \tilde{Y}_p(\mathbf{y})}{\partial x(\mathbf{y})} - \frac{\partial \tilde{B}(\mathbf{y})}{\partial x(\mathbf{y})},$$

where $\frac{\partial \tilde{Y}_p(\mathbf{y})}{\partial x(\mathbf{y})}$ and $\frac{\partial \tilde{B}(\mathbf{y})}{\partial x(\mathbf{y})}$ are the single-spell partial derivatives of $Y_p(\mathbf{y})$ and $B(\mathbf{y})$ with respect to $x(\mathbf{y})$.

Proof. Application of the Envelope Theorem to the identity $M(\mathbf{y}) \equiv L(\mathbf{y}) + tY_p(\mathbf{y}) - B(\mathbf{y})$. ■

D.1 Employment Rungs ($\mathbf{y} \neq \mathbf{y}_0$)

Consider any employment rung $\mathbf{y} \in \mathscr{Y} \setminus \{\mathbf{y}_0\}$ on the job ladder. A marginal increase in the income tax rate generates a rung-$\mathbf{y}$ deadweight loss which springs from three sources: distorted search effort, distorted reservation amenities, and distorted vacancy creation. We consider each the sources in turn.

Since we here consider $\mathbf{y} \neq \mathbf{y}_0$, (D1) implies $\tilde{B}(\mathbf{y}) = B_1$, which further implies $\frac{\partial \tilde{B}(\mathbf{y})}{\partial e(\mathbf{y})} = \frac{\partial \tilde{B}(\mathbf{y})}{\partial \phi(y_p', \mathbf{y})} = \frac{\partial \tilde{B}(\mathbf{y})}{\partial \theta(\mathbf{y})} = 0$. Hence, for $\mathbf{y} \neq \mathbf{y}_0$, there are no deadweight losses coming from aggregate unemployment expenditures. Indeed, from Lemma 3, the single spell partial deadweight losses defined in the main text are given by $O_t^e(\mathbf{y}) \equiv -\frac{\partial \tilde{M}_p(\mathbf{y})}{\partial e(\mathbf{y})} \frac{\partial e(\mathbf{y})}{\partial t} = -t \frac{\partial \tilde{Y}_p(\mathbf{y})}{\partial e(\mathbf{y})} \frac{\partial e(\mathbf{y})}{\partial t}$, $O_t^{\phi}(\mathbf{y}) \equiv -\int \frac{\partial \tilde{M}_p(\mathbf{y})}{\partial \phi(y_p', \mathbf{y})} \frac{\partial \phi(y_p', \mathbf{y})}{\partial t} dy_p' = -t \int \frac{\partial \tilde{Y}_p(\mathbf{y})}{\partial \phi(y_p', \mathbf{y})} \frac{\partial \phi(y_p', \mathbf{y})}{\partial t} dy_p'$, and $O_t^{\theta}(\mathbf{y}) \equiv -\frac{\partial \tilde{M}_p(\mathbf{y})}{\partial \theta(\mathbf{y})} \frac{\partial \theta(\mathbf{y})}{\partial t} = -t \frac{\partial \tilde{Y}_p(\mathbf{y})}{\partial \theta(\mathbf{y})} \frac{\partial \theta(\mathbf{y})}{\partial t}$.

Search effort. The equilibrium contract split the total marginal gain from rung-$\mathbf{y}$ search effort, $p(\theta(\mathbf{y}))\Delta L(\mathbf{y})$, such that the firm takes $(1-\beta)p(\theta(\mathbf{y}))\Delta L(\mathbf{y}) = (1-t)\theta(\mathbf{y})\overline{\pi}$ and the worker takes $\beta p(\theta(\mathbf{y}))\Delta L(\mathbf{y}) = p(\theta(\mathbf{y}))\Delta L(\mathbf{y}) - (1-t)\theta(\mathbf{y})\overline{\pi}$, see (17) and (18). Using $L(\mathbf{y}) = (1-t)Y_p(\mathbf{y}) + Y_z(\mathbf{y})$, (17), the first order condition for worker search effort $e(\mathbf{y})$, can therefore be expressed as

$$c'(e(\mathbf{y})) = p(\theta(\mathbf{y}))\left[(1-t)\Delta Y_p(\mathbf{y}) + \Delta Y_z(\mathbf{y})\right] - (1-t)\theta(\mathbf{y})\overline{\pi}. \tag{D2}$$

Taking the derivative of (D2) with respect to t, and subsequently using $\overline{\pi} = \frac{1-\beta}{1-t} q(\theta(\mathbf{y}))[(1-t)\Delta Y_p(\mathbf{y}) + \Delta Y_z(\mathbf{y})]$, see (18), yields

$$\frac{\partial e(\mathbf{y})}{\partial t} = -\frac{p(\theta(\mathbf{y}))}{c''(e(\mathbf{y}))}\left[\frac{\beta(1-t)\Delta Y_p(\mathbf{y}) - (1-\beta)\Delta Y_z(\mathbf{y})}{1-t}\right]. \tag{D3}$$

Changes to workers' search effort at rung-$\mathbf{y}$ impacts the rung-$\mathbf{y}$ tax base which gives rise to a single-spell partial deadweight loss of $O_t^e(\mathbf{y}) \equiv -\frac{\partial \tilde{M}_p(\mathbf{y})}{\partial e(\mathbf{y})} \frac{\partial e(\mathbf{y})}{\partial t} = -t \frac{\partial \tilde{Y}_p(\mathbf{y})}{\partial e(\mathbf{y})} \frac{\partial e(\mathbf{y})}{\partial t}$. Keep in mind that $Y_p(\mathbf{y})$ is given by (11), that $p(\theta) = \theta(\mathbf{y})q(\theta(\mathbf{y}))$, and that $\overline{\pi} = \frac{1-\beta}{1-t} q(\theta(\mathbf{y}))[(1-t)\Delta Y_p(\mathbf{y}) + \Delta Y_z(\mathbf{y})]$. Then, from (11) we have that

$$R(\mathbf{y})^{-1} \frac{\partial \tilde{Y}_p(\mathbf{y})}{\partial e(\mathbf{y})} = p(\theta(\mathbf{y}))\Delta Y_p(\mathbf{y}) - (1-t)\theta(\mathbf{y})\overline{\pi} = p(\theta(\mathbf{y}))\left[\frac{\beta(1-t)\Delta Y_p(\mathbf{y}) - (1-\beta)\Delta Y_z(\mathbf{y})}{1-t}\right], \tag{D4}$$

which, upon multiplying by $R(\mathbf{y})$, t, and (D3), yields the expression for $O_t^e(\mathbf{y})$ in (29).

Job ranking. Consider first how a marginal change in the tax rate impact $\phi(y_p', \mathbf{y})$, the reservation amenity associated with a productivity attribute draw y_p' for a worker currently employed in a rung-$\mathbf{y}$ job. It follows from Definition 2, that

$$\frac{\partial \phi(y_p', \mathbf{y})}{\partial t} = y_p' - y_p; \tag{D5}$$

Using (11), and keeping in mind that $\Delta Y_p(\mathbf{y}) = \int_{\underline{y}_p}^{\overline{y}_p} \int_{\phi(y_p', \mathbf{y})}^{\overline{y}_z} [Y_p(\mathbf{y}') - Y_p(\mathbf{y})] dF(\mathbf{y})$, we obtain

$$\begin{aligned} \frac{\partial \tilde{Y}_p(\mathbf{y})}{\partial \phi(y_p', \mathbf{y})} &= -R(\mathbf{y})e(\mathbf{y})p(\theta(\mathbf{y}))[Y_p(y_p', \phi(y_p', \mathbf{y})) - Y_p(\mathbf{y})] f_p(y_p') f_z(\phi(y_p', \mathbf{y})) \\ &= -R(\mathbf{y})^2 e(\mathbf{y})p(\theta(\mathbf{y})) f_p(y_p') f_z(\phi(y_p', \mathbf{y}))(y_p' - y_p), \end{aligned} \tag{D6}$$

because workers are indifferent between type-$\mathbf{y}' = (y_p', \phi(y_p', \mathbf{y}))$ jobs and type-$\mathbf{y}$ jobs, with the implication that the equilibrium contracts in the $\mathbf{y}'$- and $\mathbf{y}$-submarkets stipulate the same search effort, job acceptance decisions,[37] and the same labour market tightness, and because the discounted time that the worker who accepts the type-$\mathbf{y}'$ job will be in that job is $R(\mathbf{y}') = R(\mathbf{y})$, after which the equilibrium paths again coincide.

From Lemma 3, using (D5) and (D6), $O_t^{\phi}(\mathbf{y})$, the single-spell partial deadweight loss associated with distorted ranking of jobs at rung-$\mathbf{y}$, is given by

$$O_t^{\phi}(\mathbf{y}) = tR(\mathbf{y})^2 e(\mathbf{y})p(\theta(\mathbf{y})) \int_{\underline{y}_p}^{\overline{y}_p} (y_p' - y_p)^2 f_p(y_p') f_z(\phi(y_p', \mathbf{y})) dy_p', \tag{D7}$$

[37] That is, $\mathscr{Y}^a(y_p', \phi(y_p', \mathbf{y})) \setminus \mathscr{Y}^a(\mathbf{y})$ has measure zero.

where we have integrated over potential productivity attribute draws. Hence, we have derived (31).

Job creation. Inserting $\Delta L(\mathbf{y}) = (1-t)\Delta Y_p(\mathbf{y}) + \Delta Y_z(\mathbf{y})$ into the zero profit condition (18) gives

$$(1-\beta)q(\theta(\mathbf{y}))\left[(1-t)\Delta Y_p(\mathbf{y}) + \Delta Y_z(\mathbf{y})\right] - (1-t)\overline{\pi} = 0.$$

Taking the derivative with respect to t, utilizing that the Cobb-Douglas matching technology implies $\mathrm{El}_\theta q(\theta) = -\beta$, and using the equilibrium condition $\overline{\pi}(1-t) = q(\theta(\mathbf{y}))(1-\beta)\left[(1-t)\Delta Y_p + \Delta Y_z\right]$ gives

$$\frac{\partial \theta(\mathbf{y})}{\partial t} = \frac{\theta(\mathbf{y})\Delta Y_z(\mathbf{y})}{\beta(1-t)\left[(1-t)\Delta Y_p(\mathbf{y}) + \Delta Y_z(\mathbf{y})\right]}. \tag{D8}$$

Using (11), $\overline{\pi} = \frac{1-\beta}{1-t}q(\theta(\mathbf{y}))[(1-t)\Delta Y_p(\mathbf{y}) + \Delta Y_z(\mathbf{y})]$, and $p'(\theta(\mathbf{y})) = (1-\beta)q(\theta(\mathbf{y}))$ we find that

$$\frac{\partial \tilde{Y}_p(\mathbf{y})}{\partial \theta(\mathbf{y})} = -R(\mathbf{y})e(\mathbf{y})(1-\beta)q(\theta(\mathbf{y}))\frac{\Delta Y_z(\mathbf{y})}{1-t}. \tag{D9}$$

From (D9) and (D8) we obtain the expression for $O_t^\theta(\mathbf{y}) = -t\frac{\partial \tilde{Y}_p(\mathbf{y})}{\partial \theta(\mathbf{y})}\frac{\partial \theta(\mathbf{y})}{\partial t}$ given in (33).

D.2 Unemployment Rungs ($\mathbf{y} = \mathbf{y}_0$)

Next, consider the unemployment rungs on the job ladder, i.e. rung-$\mathbf{y}_0$. By assumption, unemployed workers accept any job offer, irrespective of the tax rate t; hence, there are no deadweight loss coming from distorted job rankings for unemployed workers. However, there is an additional fiscal externality that contributes to the deadweight loss in the unemployment submarkets, operating via the expenditure side of the government budget: changes to search effort and tightness in the unemployment submarket affect NPV unemployment benefits, B_0, as well as the tax base.

Indeed, since $\mathbf{y} = \mathbf{y}_0$, we have $B(\mathbf{y}_0) = B_0$, see (D1), and thus $\frac{\partial \tilde{B}(\mathbf{y}_0)}{\partial e(\mathbf{y}_0)} \neq 0$ and $\frac{\partial \tilde{B}(\mathbf{y}_0)}{\partial \theta(\mathbf{y}_0)} \neq 0$. Lemma 3 therefore implies that the single spell partial deadweight losses from unemployed search effort and vacancy creation in the unemployment submarket can be computed as $O_t^e(\mathbf{y}_0) \equiv -\frac{\partial \tilde{M}_p(\mathbf{y})}{\partial e(\mathbf{y})}\frac{\partial e(\mathbf{y})}{\partial t} = \left[-t\frac{\partial \tilde{Y}_p(\mathbf{y})}{\partial e(\mathbf{y})} + \frac{\partial \tilde{B}(\mathbf{y})}{\partial e(\mathbf{y})}\right]\frac{\partial e(\mathbf{y})}{\partial t}$ and $O_t^\theta(\mathbf{y}) \equiv -\frac{\partial \tilde{M}_p(\mathbf{y})}{\partial \theta(\mathbf{y})}\frac{\partial \theta(\mathbf{y})}{\partial t} = \left[-t\frac{\partial \tilde{Y}_p(\mathbf{y})}{\partial \theta(\mathbf{y})} + \frac{\partial \tilde{B}(\mathbf{y})}{\partial \theta(\mathbf{y})}\right]\frac{\partial \theta(\mathbf{y})}{\partial t}$.

Search effort. We can compute the (single-spell partial) impact of a marginal change to unemployed search effort on the tax base in the type-$\mathbf{y}_0$ submarket as $\frac{\partial \tilde{Y}_p(\mathbf{y}_0)}{\partial e(\mathbf{y}_0)}$. Using exactly the same argument as for employed workers, we find that the rung-$\mathbf{y}_0$ specific deadweight losses for search effort operating through the revenue channel is given by the right-hand side of (29), with $\mathbf{y}_0$ substituted for $\mathbf{y}$, and with $R(\mathbf{y})$ replaced by $R_0 \equiv [r + e(\mathbf{y}_0)p(\theta(\mathbf{y}_0)]^{-1}$, the discounted expected unemployment duration.

Turning now to the deadweight losses (or gains, as the case may be) associated with B_0, i.e. operating through the expenditure channel, the impact of a marginal distortion to job search effort $e(\mathbf{y}_0)$ on B_0 for the *current* unemployment spell (i.e. holding B_1 constant) is

$$\frac{\partial \tilde{B}_0}{\partial e(\mathbf{y}_0)} = R_0 p(\theta(\mathbf{y}_0))(B_1 - B_0) < 0, \tag{D10}$$

with associated marginal deadweight loss

$$\frac{\partial \tilde{B}_0}{\partial e(\mathbf{y}_0)}\frac{\partial e(\mathbf{y}_0)}{\partial t} = -R_0\frac{p(\theta(\mathbf{y}))^2}{c''(e(\mathbf{y}))}\left[\frac{\beta(1-t)\Delta Y_p(\mathbf{y}) - (1-\beta)\Delta Y_z(\mathbf{y})}{1-t}\right](B_1 - B_0), \tag{D11}$$

Adding up the revenue-side (single-spell partial) deadweight loss $-t\frac{\partial \tilde{Y}_p(\mathbf{y}_0)}{\partial e(\mathbf{y}_0)}\frac{\partial e(\mathbf{y}_0)}{\partial t}$ and expenditure-side (single-spell partial) deadweight loss $\frac{\partial \tilde{B}_0}{\partial e(\mathbf{y}_0)}\frac{\partial e(\mathbf{y}_0)}{\partial t}$ yields the expression for $O_t^e(\mathbf{y}_0)$ given in (35).

Vacancy creation. We compute the single-spell partial impact of a marginal change to labour market tightness in the unemployment submarket on the tax base as $\frac{\partial \tilde{Y}_p(\mathbf{y}_0)}{\partial \theta(\mathbf{y}_0)}$, which is given by the right-hand side of (33) with $\mathbf{y}_0$ substituted for $\mathbf{y}$, and with $R(\mathbf{y})$ replaced by R_0.

With respect to the deadweight losses on the expenditure-side of the government budget, the impact of a marginal distortion to labour market tightness in the type-$\mathbf{y}_0$ submarket, $\theta(\mathbf{y}_0)$, on B_0 for the current unemployment spell is

$$\frac{\partial \tilde{B}_0}{\partial \theta(\mathbf{y}_0)} = R_0 e(\mathbf{y}_0) p'(\theta(\mathbf{y}_0))(B_1 - B_0) < 0, \tag{D12}$$

with associated marginal deadweight loss

$$\frac{\partial \tilde{B}_0}{\partial \theta(\mathbf{y}_0)} \frac{\partial \theta(\mathbf{y}_0)}{\partial t} = R_0 \frac{(1-\beta) e(\mathbf{y}_0) p(\theta(\mathbf{y})) \Delta Y_z(\mathbf{y})}{\beta(1-t)\left[(1-t)\Delta Y_p(\mathbf{y}) + \Delta Y_z(\mathbf{y})\right]}(B_1 - B_0). \tag{D13}$$

Adding up the revenue-side (single-spell partial) deadweight loss $-t\frac{\partial \tilde{Y}_p(\mathbf{y}_0)}{\partial \theta(\mathbf{y}_0)}\frac{\partial \theta(\mathbf{y}_0)}{\partial t}$ and expenditure-side (single-spell partial) deadweight loss $\frac{\partial \tilde{B}_0}{\partial \theta(\mathbf{y}_0)}\frac{\partial \theta(\mathbf{y}_0)}{\partial t}$ yields the expression for $O_t^\theta(\mathbf{y}_0)$ given in (37).

E Aggregating $O_t^e(\mathbf{y})$, $O_t^\phi(\mathbf{y})$ and $O_t^\theta(\mathbf{y})$

To avoid unecessarily complicated notation, Appendix E ignores worker-type heterogeneity.

To characterize the constituent components of the marginal deadweight loss decomposition (38), starting with $M_t^e(\mathbf{y}_0)$, take the derivative of $M(\mathbf{y}_0)$ with respect to $e(\mathbf{y}')$, multiply by $\frac{\partial e(\mathbf{y}')}{dt}$, and integrate with respect to $\mathbf{y}' \in \mathscr{Y}$. This yields

$$M_t^e(\mathbf{y}_0) = O_t^e(\mathbf{y}_0) + \int_{\underline{y}_p}^{\overline{y}_p}\int_{\underline{y}_z}^{\overline{y}_z} \xi(\mathbf{y}_0, \mathbf{y}) M_t^e(\mathbf{y}) d\mathbf{y}, \tag{E1}$$

where

$$\xi(\mathbf{y}_0, \mathbf{y}) = \frac{e(\mathbf{y}_0) p(\theta(\mathbf{y}_0)) f(\mathbf{y}')}{r + e(\mathbf{y}_0) p(\theta(\mathbf{y}_0))} \tag{E2}$$

is the discounted density that an unemployment spell terminates with the worker moving to a rung-$\mathbf{y}$ job and $M_t^e(\mathbf{y}) \equiv \int_{\underline{y}_p}^{\overline{y}_p}\int_{\underline{y}_z}^{\overline{y}_z} \frac{\partial M(\mathbf{y})}{\partial e(\mathbf{y}')}\frac{de(\mathbf{y}')}{dt} d\mathbf{y}'$ is the part of the impact of a marginal change to income taxation on NPV match utility in a rung-$\mathbf{y}$ match that derives from distorted job search effort.

However, because a rung-$\mathbf{y}$ worker may transition to unemployment, $M_t^e(\mathbf{y})$ encompasses $M_t^e(\mathbf{y}_0)$. To see this, define $\hat{M}(\mathbf{y}) \equiv M(\mathbf{y}) - \frac{s}{r+s}M(\mathbf{y}_0)$ to be the expected before-tax NPV match utility of an *employment cycle with starting-rung* $\mathbf{y}$: a sequence of jobs uninterrupted by unemployment, starting with a rung-$\mathbf{y}$ job. From (10) we have that

$$R(\mathbf{y})^{-1}\hat{M}(\mathbf{y}) = y_p + y_z - c(e(\mathbf{y})) + p(\theta(\mathbf{y}))e(\mathbf{y}) \int_{\underline{y}_p}^{\overline{y}_p}\int_{\phi(y_p', \mathbf{y})}^{\overline{y}_z} \hat{M}(\mathbf{y}')dF(\mathbf{y}') - e(\mathbf{y})\theta(\mathbf{y})\overline{\pi}, \tag{E3}$$

where $R(\mathbf{y}) \equiv [r+s+e(\mathbf{y})p(\theta(\mathbf{y}))\Pr(\mathbf{y} \in \mathscr{Y}_a(\mathbf{y}))]^{-1}$ is the discounted expected duration of stay at rung-$\mathbf{y}$. By construction, $\hat{M}(\mathbf{y})$ is independent of $M(\mathbf{y}_0)$; moreover, with $\hat{M}_t^e(\mathbf{y}) \equiv \int_{\underline{y}_p}^{\overline{y}_p}\int_{\underline{y}_z}^{\overline{y}_z} \frac{\partial \hat{M}(\mathbf{y})}{\partial e(\mathbf{y}')}\frac{de(\mathbf{y}')}{dt} d\mathbf{y}'$, $M_t^e(\mathbf{y}) = \hat{M}_t^e(\mathbf{y}) + \frac{s}{r+s}M_t^e(\mathbf{y}_0)$. Substituting this expression for $M_t^e(\mathbf{y})$ into (E1) yields:

$$M_t^e(\mathbf{y}_0) = \frac{r+s}{r}\left[\frac{r + e(\mathbf{y}_0)p(\theta(\mathbf{y}_0))}{r + s + e(\mathbf{y}_0)p(\theta(\mathbf{y}_0))}\right]\left[O_t^e(\mathbf{y}_0) + \int_{\underline{y}_p}^{\overline{y}_p}\int_{\underline{y}_z}^{\overline{y}_z} \xi(\mathbf{y}_0, \mathbf{y})\hat{M}_t^e(\mathbf{y}) d\mathbf{y}\right]. \tag{E4}$$

Corresponding expressions of $M_t^\phi(\mathbf{y}_0)$ and $M_t^\theta(\mathbf{y}_0)$ in terms of $\hat{M}_t^\phi(\mathbf{y})$ and $\hat{M}_t^\theta(\mathbf{y})$ have the same structure and obtains by replacing e by ϕ and θ in (E4).

The recursive representation of $\hat{M}(\mathbf{y})$ in (E3) maps any $\mathbf{y} \in \mathscr{Y} \setminus \{\mathbf{y}_0\}$ to the before-tax NPV utility of an employment cycle starting at that rung-$\mathbf{y}$ via the policy functions e, ϕ, and θ. That is, $\hat{M}(\mathbf{y})$ depends on search effort, reservation amenities, and tightness at each employment rung of the job ladder that the worker may visit during an employment cycle starting at rung-$\mathbf{y}$. Therefore, the derivative of $\hat{M}(\mathbf{y})$ with respect to the income tax rate t is given by

$$\frac{d\hat{M}(\mathbf{y})}{dt} \equiv \hat{M}_t^e(\mathbf{y}) + \hat{M}_t^\phi(\mathbf{y}) + \hat{M}_t^\theta(\mathbf{y}), \tag{E5}$$

where $\hat{M}_t^e(\mathbf{y}) \equiv \int_{\mathscr{Y}\setminus\{\mathbf{y}_0\}} \frac{\partial \hat{M}(\mathbf{y})}{\partial e(\mathbf{y}')} \frac{de(\mathbf{y}')}{dt} d\mathbf{y}'$, $\hat{M}_t^\phi(\mathbf{y}) \equiv \int_{\mathscr{Y}\setminus\{\mathbf{y}_0\}} \int_{\mathscr{Y}_p\setminus\{b\}} \frac{\partial \hat{M}(\mathbf{y})}{\partial \phi(y_p'',\mathbf{y}')} \frac{d\phi(y_p'',\mathbf{y}')}{dt} dy_p'' d\mathbf{y}'$, and $\hat{M}_t^\theta(\mathbf{y}) \equiv \int_{\mathscr{Y}\setminus\{\mathbf{y}_0\}} \frac{\partial \hat{M}(\mathbf{y})}{\partial \theta(\mathbf{y}')} \frac{d\theta(\mathbf{y}')}{dt} d\mathbf{y}'$. Equation (E5) decomposes the marginal deadweight loss from income taxation in an employment cycle with starting rung-$\mathbf{y}$ into deadweight losses $\hat{M}_t^e(\mathbf{y})$, $\hat{M}_t^\phi(\mathbf{y})$, and $\hat{M}_t^\theta(\mathbf{y})$ stemming from distortions to job search effort, job rankings, and vacancy creation.

The remainder of this appendix uses the job ladder structure to express $\hat{M}_t^e(\mathbf{y})$, $\hat{M}_t^\phi(\mathbf{y})$, and $\hat{M}_t^\theta(\mathbf{y})$ as integrals of the single-spell partial deadweight losses, $O_t^{e(\mathbf{y})}(\mathbf{y})$, $O_t^{\phi(y_p',\mathbf{y})}(\mathbf{y})$, and $O_t^{\theta(\mathbf{y})}(\mathbf{y})$, discussed in Section 5. This will yield equation (39) in the main text.

Consider first $\hat{M}_t^e$. Using (E3), take the derivative of $\hat{M}(\mathbf{y})$ with respect to generic rung-$\mathbf{y}''$ search effort, $e(\mathbf{y}'')$, multiply by $\frac{de(\mathbf{y}'')}{dt}$, and integrate with respect to $\mathbf{y}'' \in \mathscr{Y}$ to obtain

$$\hat{M}_t^e(\mathbf{y}) = -O_t^e(\mathbf{y}) + \int_{\underline{y}_p}^{\overline{y}_p} \int_{\phi(y_p',\mathbf{y})}^{\overline{y}_z} \xi(\mathbf{y},\mathbf{y}') \hat{M}_t^e(\mathbf{y}') d\mathbf{y}', \tag{E6}$$

where $O_t^e(\mathbf{y}) \equiv -\frac{\partial \hat{M}(\mathbf{y})}{\partial e(\mathbf{y})} \frac{de(\mathbf{y})}{dt} = -t\frac{\partial M(\mathbf{y})}{\partial e(\mathbf{y})} \frac{de(\mathbf{y})}{dt} = -t\frac{\partial W(\mathbf{y})}{\partial e(\mathbf{y})} \frac{de(\mathbf{y})}{dt}$ is the own-rung-$\mathbf{y}$ deadweight loss stemming from distortions to job search effort at rung-$\mathbf{y}$, derived in Section 5 of the main text, and where

$$\xi(\mathbf{y},\mathbf{y}') = \frac{e(\mathbf{y})p(\theta(\mathbf{y}))f(\mathbf{y}')}{r + s + e(\mathbf{y})p(\theta(\mathbf{y}))\Pr(\mathbf{y}' \in \mathscr{Y}^a(\mathbf{y}))} \tag{E7}$$

is the discounted density of a rung-$\mathbf{y}$ job terminating with the worker moving to a rung-$\mathbf{y}'$ job.[38]

The right-hand side of equation (E6) defines a linear map $\Gamma_{O_t^e}$ from the set of continuous functions on $\mathscr{Y}$, denoted $\mathscr{C}(\mathscr{Y})$, onto itself; that is, for any $x \in \mathscr{C}(\mathscr{Y})$, $\Gamma_{O_t^e} : x(\mathbf{y}) \mapsto -O_t^e(\mathbf{y}) + \int_{\underline{y}_p}^{\overline{y}_p} \int_{\phi(y_p',\mathbf{y})}^{\overline{y}_z} \xi(\mathbf{y},\mathbf{y}')x(\mathbf{y}')d\mathbf{y}'$. The function $\hat{M}_t^e$, is a fixed point of $\Gamma_{O_t^e}$.

Repeating the calculations for reservation amenities and labour market tightness shows that the functions $\hat{M}_t^\phi$ and $\hat{M}_t^\theta$ are implicitly defined as fixed points of mappings like (E6), only with O_t^e replaced by the own-rung marginal deadweight losses associated with reservation amenities O_t^ϕ and labour market tightness O_t^θ, respectively.[39]

In summary, for $k \in \{e, \phi, \theta\}$, define $\Gamma_{O_t^k} : \mathscr{C}(\mathscr{Y}) \to \mathscr{C}(\mathscr{Y})$ such that

$$\Gamma_{O_t^k} : x(\mathbf{y}) \mapsto O_t^k(\mathbf{y}) + \int_{\underline{y}_p}^{\overline{y}_p} \int_{\phi(y_p',\mathbf{y})}^{\overline{y}_z} \xi(\mathbf{y},\mathbf{y}')x(\mathbf{y}')d\mathbf{y}'. \tag{E8}$$

The functions $\hat{M}_t^e$, $\hat{M}_t^\phi$, and $\hat{M}_t^\theta$, are fixed points of $\Gamma_{O_t^e}$, $\Gamma_{O_t^\phi}$, and $\Gamma_{O_t^\theta}$, respectively.

---

[38] To see this, write

$$\xi(\mathbf{y},\mathbf{y}') = \left[\frac{s + e(\mathbf{y})p(\theta(\mathbf{y}))\Pr(\mathbf{y}' \in \mathscr{Y}^a(\mathbf{y}))}{r + s + p(\theta(\mathbf{y}))e(\mathbf{y})\Pr(\mathbf{y}' \in \mathscr{Y}^a(\mathbf{y}))}\right]\left[\frac{e(\mathbf{y})p(\theta(\mathbf{y}))f(\mathbf{y}')}{s + e(\mathbf{y})p(\theta(\mathbf{y}))\Pr(\mathbf{y}' \in \mathscr{Y}^a(\mathbf{y}))}\right].$$

The second term in the right-hand side product is the density that a rung-$\mathbf{y}$ spell ends with the worker making a transition to a rung-$\mathbf{y}'$ job. The first term is the discount factor that must be applied to a payment expected to received $1/[s+e(\mathbf{y})p(\theta(\mathbf{y}))\Pr(\mathbf{y}' \in \mathscr{Y}^a(\mathbf{y}))]$ units of time into the future, when—as is the case in our model—the duration until the payment is made follows an Exponential distribution.

[39] For reservation amenities, the exercise entails taking the derivative of (E3) with respect to $\phi(y_p', \mathbf{y}'')$, multiplying by $\frac{d\phi(y_p',\mathbf{y}'')}{dt}$, integrating with respect to $y_p' \in \mathscr{Y}_p$, and then integrating with respect to $\mathbf{y}'' \in \mathscr{Y}$.

Lemma 4 *Let $\mathscr{C}(\mathscr{Y})$ be the space of continuous functions on $\mathscr{Y}$ with the sup norm. Then, for $k \in \{e, \phi, \theta\}$, any $O_t^k : \mathscr{Y} \setminus \{\mathbf{y}_0\} \to \mathbb{R}$, and $x \in \mathscr{C}(\mathscr{Y})$, the linear map $\Gamma_{O_t^k} : x(\mathbf{y}) \mapsto O_t^k(\mathbf{y}) + \int_{\underline{y}_p}^{\bar{y}_p} \int_{\phi(y_p',\mathbf{y})}^{\bar{y}_z} \xi(\mathbf{y}, \mathbf{y}')x(\mathbf{y}')d\mathbf{y}'$ is a contraction.*

Proof. We apply Blackwell's sufficient conditions for a contraction (see e.g. Stokey and Lucas, 1996, Theorem 3.3). As $\mathscr{Y}$ is compact, elements in $\mathscr{C}(\mathscr{Y})$ are bounded functions. We then need to verify that $\Gamma_{O_t^k}$ satisfies monotonicity and discounting.

Clearly, $\Gamma_{O_t^k}$ is monotone. To verify that $\Gamma_{O_t^k}$ satisfies discounting, let $\bar{y}_p$ and $\bar{y}_z$ denote the upper bounds of the supports of y_p and y_z, respectively. An upper bound on $\hat{M}$ (and M) is then $\bar{M} = [(1-t)\bar{y}_p + \bar{y}_z]/r$, such that upper bounds on e and θ, denoted $\bar{e}$ and $\bar{\theta}$, respectively, are implicitly given by $c'(\bar{e}) = p(\bar{\theta})\bar{M}$ and $\bar{\pi} = q(\bar{\theta})\bar{M}$, respectively. Since $\xi(\mathbf{y}, \mathbf{y}')$ is increasing in $e(\mathbf{y})$ and $\theta(\mathbf{y})$, it follows that $\xi(\mathbf{y}, \mathbf{y}') \leq \bar{e}p(\bar{\theta})f(\mathbf{y}')/[r + s + \bar{e}p(\bar{\theta}) \Pr(\mathbf{y}' \in \mathscr{Y}^a(\mathbf{y}))]$, which, provided $r + s > 0$, implies that

$$\int_{\underline{y}_p}^{\bar{y}_p} \int_{\phi(y_p',\mathbf{y})}^{\bar{y}_z} \xi(\mathbf{y}, \mathbf{y}')d\mathbf{y}' \leq \frac{\bar{e}p(\bar{\theta}) \Pr(\mathbf{y}' \in \mathscr{Y}^a(\mathbf{y}))}{r + s + \bar{e}p(\bar{\theta}) \Pr(\mathbf{y}' \in \mathscr{Y}^a(\mathbf{y}))} \leq \frac{\bar{e}p(\bar{\theta})}{r + s + \bar{e}p(\bar{\theta})} \in (0, 1).$$

Now, take any $x \in \mathscr{C}(\mathscr{Y})$. Then,

$$[\Gamma_{O_t^k}(x + a)](\mathbf{y}) = [\Gamma_{O_t^k}x](\mathbf{y}) + a\int_{\underline{y}_p}^{\bar{y}_p} \int_{\phi(y_p',\mathbf{y})}^{\bar{y}_z} \xi(\mathbf{y}, \mathbf{y}')d\mathbf{y} \leq [\Gamma_{O_t^k}z](\mathbf{y}) + a\frac{p(\bar{\theta})\bar{e}}{r + s + p(\bar{\theta})\bar{e}}.$$

Hence, there exist some $\beta \in (0, 1)$—indeed, take any $\beta \in [p(\bar{\theta})\bar{e}/[r + s + p(\bar{\theta})\bar{e}], 1)$—such that $[\Gamma_{O_t^k}(z + a)](\mathbf{y}) \leq [\Gamma_{O_t^k}z](\mathbf{y}) + \beta a$. That is, $\Gamma_{O_t^k}$ satisfies discounting. ■

Lemma 4 states that $\Gamma_{O_t^k}$ is a contraction mapping. The Contraction Mapping Theorem (see e.g. Stokey and Lucas, 1996, Theorem 3.2) then ensures that $\Gamma_{O_t^e}$, $\Gamma_{O_t^\phi}$ and $\Gamma_{O_t^\theta}$ each has a unique fixed point. That is, the derivative functions $\hat{M}_t^e$, $\hat{M}_t^\phi$ and $\hat{M}_t^\theta$ exist and are unique.

Next, we show that $\hat{M}_t^e$, $\hat{M}_t^\phi$ and $\hat{M}_t^\theta$ can be expressed as weighted averages of the own-rung marginal deadweight losses, O_t^e, O_t^ϕ and O_t^θ. As $\hat{M}_t^e$, $\hat{M}_t^\phi$ and $\hat{M}_t^\theta$ exist and are unique for any O_t^k, (E8) implies that

$$O_t^k(\mathbf{y}) = \hat{M}_t^k(\mathbf{y}) - \Gamma_0\hat{M}_t^k(\mathbf{y}) = [1 - \Gamma_0]\hat{M}_t^k(\mathbf{y}), \tag{E9}$$

where Γ_0 is given by (E8) with $O_t^k = 0$. We will show that the linear map $[1 - \Gamma_0]$ is injective, and hence invertible.[40] To that end, we first define the null-space of a linear map, and show that a linear map is injective if and only if the only element in its null-space is the null-vector.

Definition 4 (Null-space) *Let $\mathscr{V}$ and $\mathscr{W}$ be vector spaces, and let $T : \mathscr{V} \to \mathscr{W}$ be a linear map. The null-space of T is the set of elements in $\mathscr{V}$ that map to the null-vector $\mathbf{0}$. That is, $\text{null}T \equiv \{\mathbf{v} \in \mathscr{V} : T(\mathbf{v}) = \mathbf{0}\}$.*

Lemma 5 *Let $\mathscr{V}$ and $\mathscr{W}$ be vector spaces, and let $T : \mathscr{V} \to \mathscr{W}$ be a linear map. Then T is injective if and only if $\text{null}T = \{\mathbf{0}\}$.*

Proof. By definition, $\mathbf{0} \in \mathscr{V}$ and $\mathbf{0} \in \text{null}T$. Suppose T is injective and take any $\mathbf{v} \in \text{null}T$. Then $T(\mathbf{v}) = \mathbf{0} = T(\mathbf{0})$. Since T is injective, this implies $\mathbf{v} = \mathbf{0}$. Now suppose $\text{null}T = \{\mathbf{0}\}$. Take $\mathbf{v}, \mathbf{v}' \in \mathscr{V}$ such that $T(\mathbf{v}) = T(\mathbf{v}')$. This implies $\mathbf{0} = T(\mathbf{v}) - T(\mathbf{v}') = T(\mathbf{v} - \mathbf{v}')$ which means that $\mathbf{v} - \mathbf{v}' \in \text{null}T$. Hence, $\mathbf{v} = \mathbf{v}'$, which shows that T is injective. ■

Lemma 6 $[1 - \Gamma_0] : \mathscr{C}(\mathscr{Y}) \to \mathscr{C}(\mathscr{Y})$ *with* $\Gamma_0 : x(\mathbf{y}) \mapsto \int_{\underline{y}_p}^{\bar{y}_p} \int_{\phi(y_p',\mathbf{y})}^{\bar{y}_z} \xi(\mathbf{y}, \mathbf{y}')x(\mathbf{y}')d\mathbf{y}'$ *is injective.*

Proof. Given Lemma 6 it is sufficient to show that $\text{null}[1 - \Gamma_0] = \{\mathbf{0}\}$, where $\mathbf{0}$ is the null-vector. Clearly, $[1 - \Gamma_0](\mathbf{0}) = \mathbf{0}$; hence, $\mathbf{0} \in \text{null}[1 - \Gamma_0]$. Next, take any $x \in \text{null}[1 - \Gamma_0]$; that is, take any x

[40] Technically, the map $[1 - \Gamma_0]$ has a unique inverse if and only if $[1 - \Gamma_0]$ is bijective, i.e. injective and surjective. Surjectivity of Γ_0 can be ensured by appropriately defining the image of Γ_0.

such that $[1-\Gamma_0](x) = \mathbf{0}$. As $[1-\Gamma_0](x) = x - \Gamma_0 x$, any element in null$[1-\Gamma_0]$ satisfies $x = \Gamma_0 x$. However, Lemma 4 states that $\Gamma_{O_t^k}$ is a contraction mapping for any O_t^k, including $O_t^k = 0$; hence, Γ_0 is a contraction mapping, and thus, have a unique fixed point, which must be $\mathbf{0}$. It follows that null$[1-\Gamma_0] = \{\mathbf{0}\}$, which proves the lemma. ■

By Lemma 6, the map $[1-\Gamma_0]$ is injective, and thus has a unique inverse, which we denote $[1-\Gamma_0]^{-1}$. It then follows from (E9) that, for $k \in \{e, \phi, \theta\}$,

$$\hat{M}_t^k = [1-\Gamma_0]^{-1} O_t^k, \tag{E10}$$

where O_t^k is the own-rung marginal deadweight loss as a function of $\mathbf{y} \in \mathscr{Y} \setminus \{\mathbf{y}_0\}$.

Equation (E10) shows that the derivative functions of interest $\hat{M}_t^k$ can be expressed as weighted averages of the single-spell partial marginal deadweight losses $O_t^k(\cdot)$: $\hat{M}_t^k(\mathbf{y}) = \int_{\underline{y}_p}^{\overline{y}_p} \int_{\underline{y}_z}^{\overline{y}_z} \xi^*(\mathbf{y}, \mathbf{y}') O_t^k(\mathbf{y}') dy'$ where $\xi^*(\mathbf{y}, \mathbf{y}')$ is a weight function. It follows that

$$M_t^k(\mathbf{y}_0) = \frac{r+s}{r} \left[\frac{r + e(\mathbf{y}_0)p(\theta(\mathbf{y}_0))}{r + s + e(\mathbf{y}_0)p(\theta(\mathbf{y}_0))} \right] \left[O_t^k(\mathbf{y}_0) + \int_{\underline{y}_p}^{\overline{y}_p} \int_{\underline{y}_z}^{\overline{y}_z} \xi(\mathbf{y}_0, \mathbf{y}) \int_{\underline{y}_p}^{\overline{y}_p} \int_{\underline{y}_z}^{\overline{y}_z} \xi^*(\mathbf{y}, \mathbf{y}') O_t^k(\mathbf{y}') d\mathbf{y}' d\mathbf{y} \right], \tag{E11}$$

from which (39), and the definitions of ω_0 and $\omega_1(\mathbf{y}')$ in the main text follows.

The analysis conducted thusfar invokes functional analysis, which makes it difficult to gauge the precise structure of the weight function $\xi^*(\mathbf{y}, \mathbf{y}')$. We can, however, glean relevant insights from a version of the model with a finite job ladder, where we can apply standard linear algebra to obtain explicit expressions for $\xi^*(\mathbf{y}, \mathbf{y}')$ that will aid interpretation.

E.1 A Job Ladder with Three Employment Rungs

We provide algebraic details on the job ladder transformation for the four-rung job ladder economy. In doing so, we focus exclusive on the job search effort compoenent $\hat{M}_t^e$.

The four rungs on the job ladder are unemployment $\mathbf{y}_0$, and three employment rungs: $\mathbf{y}_1$, $\mathbf{y}_2$, and $\mathbf{y}_3$. Workers rank the the four rungs as follows: $\mathbf{y}_0 \prec \mathbf{y}_1 \prec \mathbf{y}_2 \prec \mathbf{y}_3$. There are seven possible employment cycles, represented by the following ordered lists: $\{\mathbf{y}_1\}$, $\{\mathbf{y}_2\}$, $\{\mathbf{y}_3\}$, $\{\mathbf{y}_1, \mathbf{y}_2\}$, $\{\mathbf{y}_1, \mathbf{y}_3\}$,$\{\mathbf{y}_2, \mathbf{y}_3\}$, $\{\mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3\}$. For $m, n > 0$, define $\xi(\mathbf{y}_m, \mathbf{y}_n)$ to be the time discounted probability that a job at rung-$\mathbf{y}_m$ ends with a job-to-job transition to a rung-$\mathbf{y}_n$ job:

$$\xi(\mathbf{y}_m, \mathbf{y}_n) = \frac{e(\mathbf{y}_m)p(\theta(\mathbf{y}_m))f(\mathbf{y}_n)}{r + s + e(\mathbf{y}_m)p(\theta(\mathbf{y}_m)) \Pr(\mathbf{y}_n \in \mathscr{Y}_a(\mathbf{y}_m))}. \tag{E12}$$

Further, upon defining

$$\hat{\mathbf{M}}_t^e \equiv \begin{bmatrix} \hat{M}_t^e(\mathbf{y}_1) \\ \hat{M}_t^e(\mathbf{y}_2) \\ \hat{M}_t^e(\mathbf{y}_3) \end{bmatrix}, \quad \mathbf{O}_t^e \equiv \begin{bmatrix} O_t^e(\mathbf{y}_1) \\ O_t^e(\mathbf{y}_2) \\ O_t^e(\mathbf{y}_3) \end{bmatrix}, \quad \text{and} \quad \Gamma_0 \equiv \begin{bmatrix} 0 & \xi(\mathbf{y}_1, \mathbf{y}_2) & \xi(\mathbf{y}_1, \mathbf{y}_3) \\ 0 & 0 & \xi(\mathbf{y}_2, \mathbf{y}_3) \\ 0 & 0 & 0 \end{bmatrix}, \tag{E13}$$

we can represent (E9) as $\mathbf{O}_t^e = \hat{\mathbf{M}}_t^e - \mathbf{\Gamma}_0 \hat{\mathbf{M}}_t^e$, and thus (E10) as

$$\hat{\mathbf{M}}_t^e = [\mathbf{I} - \mathbf{\Gamma}_0]^{-1} \mathbf{O}_t^e, \tag{E14}$$

where

$$\mathbf{I} - \mathbf{\Gamma}_0 \equiv \begin{bmatrix} 1 & -\xi(\mathbf{y}_1, \mathbf{y}_2) & -\xi(\mathbf{y}_1, \mathbf{y}_3) \\ 0 & 1 & -\xi(\mathbf{y}_2, \mathbf{y}_3) \\ 0 & 0 & 1 \end{bmatrix}, \tag{E15}$$

and where $\mathbf{I}$ is the identity matrix.

Clearly, the upper-triangular matrix $\mathbf{I} - \mathbf{\Gamma}_0$ has determinant $|\mathbf{I} - \mathbf{\Gamma}_0| = 1$. Hence, $\mathbf{I} - \mathbf{\Gamma}_0$ is invertible with inverse matrix

$$[\mathbf{I}-\mathbf{\Gamma}_0]^{-1} \equiv \begin{bmatrix} 1 & \xi^*(\mathbf{y}_1,\mathbf{y}_2) & \xi^*(\mathbf{y}_1,\mathbf{y}_3) \\ 0 & 1 & \xi^*(\mathbf{y}_2,\mathbf{y}_3) \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & \xi(\mathbf{y}_1,\mathbf{y}_2) & \xi(\mathbf{y}_1,\mathbf{y}_2)\xi(\mathbf{y}_2,\mathbf{y}_3)+\xi(\mathbf{y}_1,\mathbf{y}_3) \\ 0 & 1 & \xi(\mathbf{y}_2,\mathbf{y}_3) \\ 0 & 0 & 1 \end{bmatrix}. \quad \text{(E16)}$$

The upper triangular entries in $[\mathbf{I}-\mathbf{\Gamma}_0]^{-1}$ constitute the $\xi^*(\mathbf{y},\mathbf{y}')$ weight function referenced above. We note that the relevant entries in the first row of $[\mathbf{I}-\mathbf{\Gamma}_0]^{-1}$ are the time-discounted probabilities that an employment cycle with starting rung $\mathbf{y}_1$ includes a spell at rung-$\mathbf{y}_1$ (unit probability, no discounting), a spell at rung-$\mathbf{y}_2$ (discounted probability $\xi(\mathbf{y}_1,\mathbf{y}_2)$), and a spell at rung-$\mathbf{y}_3$ (discounted probability $\xi(\mathbf{y}_1,\mathbf{y}_2)\xi(\mathbf{y}_2,\mathbf{y}_3)+\xi(\mathbf{y}_1,\mathbf{y}_3)$ because the worker may transit from rung-$\mathbf{y}_1$ to rung-$\mathbf{y}_3$ via a spell at rung-$\mathbf{y}_2$, an event that occurs with discounted probability $\xi(\mathbf{y}_1,\mathbf{y}_2)\xi(\mathbf{y}_2,\mathbf{y}_3)$, or the worker may transit directly from rung-$\mathbf{y}_1$ to rung-$\mathbf{y}_3$, an event that occurs with discounted probability $\xi(\mathbf{y}_1,\mathbf{y}_3)$). Similar interpretations can be given to the relevant entries in the second row, which pertains to employment cycles with starting-rung $\mathbf{y}_2$, and the third row, pertaining employment cycles with starting rung $\mathbf{y}_3$.

We surmise that these insights generalize to the theoretically appealing case of a continuous job ladder, and we therefore state in the main text that the weight function $\xi^*(\mathbf{y},\mathbf{y}')$ in (39) encompasses the likelihood that an employment cycle starting at rung-$\mathbf{y}$ involves a spell at rung-$\mathbf{y}'$.

Appendix F.3 presents the single-spell partial deadweight losses and the weights from (39) for our calibrated economy.

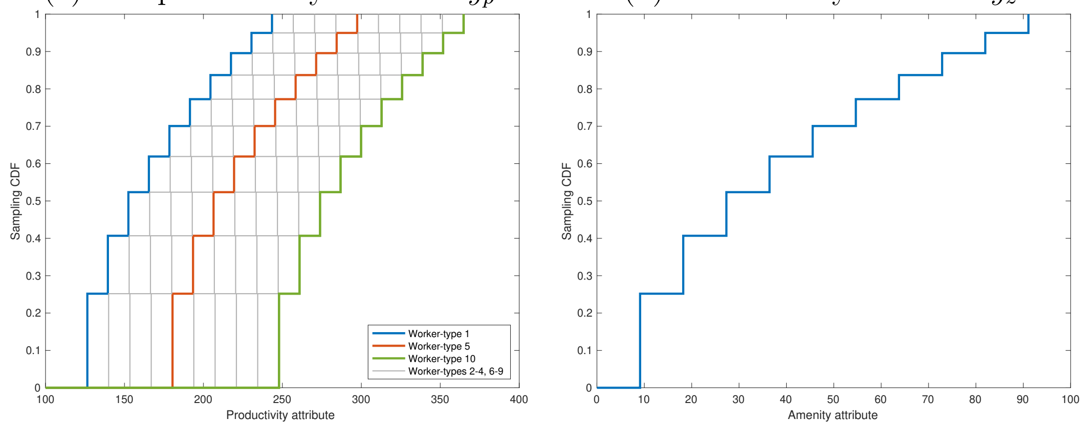
F The Calibrated Economy

F.1 The Sampling Distribution

The sampling distributions of productivity-attributes y_p and amenity-attributes y_z implied by the calibrated loading coefficients $\bar{z}$, ϱ_1, and ϱ_2 from Table 2, and the distributions of a, p, and z in Figure 4, are plotted in Figure F.1. Panel (a) plots the sampling distributions of productivity attributes which encode the fundamental source of worker-heterogeneity in the model: high-type workers face more favourable sampling distributions of productivity-attributes than low-type workers. All worker-types face the same amenity-attribute sampling distribution, see panel (b).

Figure F.1: Sampling Distributions of Productivity- and Amenity-Attributes

(a) The productivity-attribute y_p (b) The amenity-attribute y_z

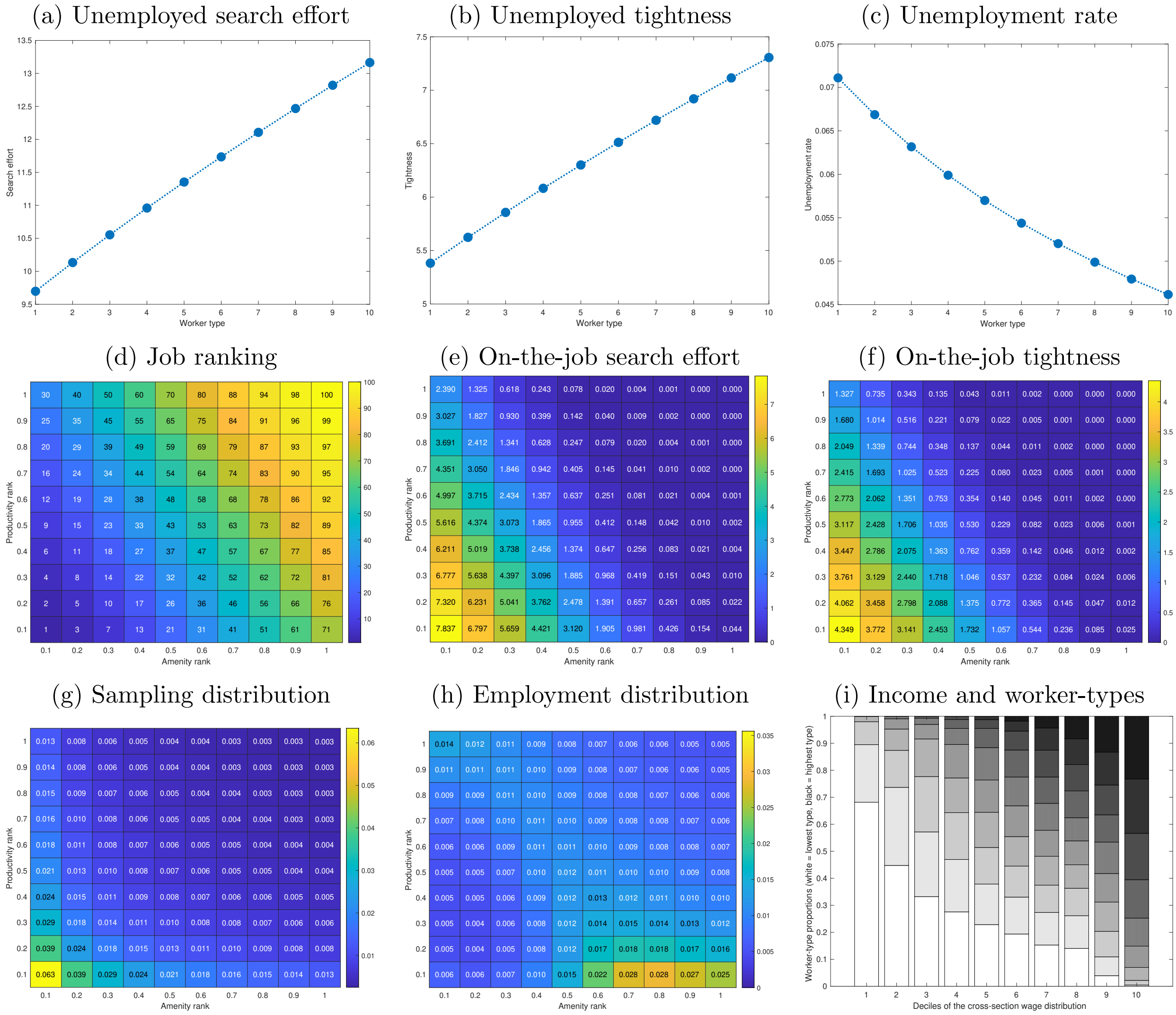


Notes: The productivity-attribute sampling distributions differ by worker-type, while all worker-types share the same amenity-attribute sampling distribution.

F.2 The Equilibrium

Figure F.2 shows key features of the calibrated equilibrium. Panels (a) and (b) plot search effort and labour market tightness by worker-type in the unemployment-submarkets. High worker-types have have higher gains from unemployed job search (see also Figure F.1). This translates into higher job search effort and tighter labour (sub)markets for higher-type workers, and an unemployment rate that declines in worker-type, as confirmed in panel (c).

Figure F.2: Calibrated Equilibrium



Panels (d), (e) and (f) in Figure F.2 concerns employed workers, and illustrates workers' ranking of job-types, search effort in each job-type, and labour market tightness in each (employed) submarket. The $\varrho_3 = 0$ restriction in (41) implies that the sampling distribution of job-attributes of different worker-types are horizontal translations of each other, with sampling distributions of higher worker-types further to the right (see Figure F.1). Once workers are employed, the expected gains from climbing the job ladder are therefore the same for all worker-types, which in turn implies that all worker-types share the same ranking, exert the same job search effort, and face the same labour market tightnesses; that is, there are no relevant dimension of worker-heterogeneity in panels (d), (e) and (f).

Panel (d) illustrates workers' ranking of jobs by productivity- and amenity-attribute in a heatmap where colder colours indicate lower ranked jobs, and warmer colours indicate higher ranked jobs, with workers' actual rank (rank 1 being the lowest, rank 100 the highest) superimposed. Naturally, workers rank the job with the lowest productivity-attribute *and* the lowest amenity-attribute the lowest, while the job with the highest productivity-attribute *and* the highest amenity-attribute is ranked the highest. In-between, workers tend to favour jobs with higher amenity and lower productivity over jobs with lower amenity and higher productivity; that is, warmer colours prevail towards the south-east corner of the heatmap and colder colours prevail towards the north-west corner.

Using similar heatmaps, panels (e) and (f) in Figure F.2 illustrate workers' job search effort and the labour market tightness they face by submarket, i.e. by productivity- and amenity-attribute, where warmer colours indicate higher values. Workers exert less search effort, and labour market tightness is lower, in higher ranked submarkets (with zero search effort and zero tightness in the highest ranked submarket), meaning that search effort and tightness are higher in the north-west corner of the heatmaps in panels (e) and (f), populated by high-productivity, low-amenity jobs, than it is in the south-west corners, which are populated by low-productivity, high-amenity jobs.

Panel (g) plots a productivity-amenity rank heatmap of the distribution from which searching workers sample job-types, i.e. F. Common distributions of p and z, see Figure 4, and the assumed independence between sampled productivity- and amenity-attributes implies that sampling distribution is symmetric around the antidiagonal running from the south-west to the north-east corner. The shape of the distribution is such that workers are more likely to be offered jobs with low productivity- and amenity-attributes than the higher-ranked jobs with high productivity- and amenity-attributes. Panel (h) plots a productivity-amenity rank heatmap of the cross section distribution of employment. Unlike the sampling distribution, the employment distribution is not symmetric because workers tend to gravitate towards the relatively higher-ranked jobs with high amenity-attributes as a result of their acceptance decisions, search effort choices, and firms' vacancy creation decisions. Indeed, the mode of the employment distribution appears in the submarket ranked 51, which has the lowest productivity-attribute, but the third-highest amenity attribute.

Finally, panel (i) shows the distribution of worker-types within each decile of the cross sectional wage distribution. While worker type and income are clearly positively correlated, it is equally clear that income is not a perfect predictor of a workers' type in the calibrated job ladder model. For example, less than 50 percent of workers of the highest type is in the highest income decile. This feature of the model limits how much redistribution across worker-types that can be achieved by an affine labour income tax system.

F.3 Details on the Deadweight Loss Decomposition

Table 3 presents the total marginal deadweight loss in the calibrated economy and decomposes it into its three constituent components: distorted job search, job ranking, and vacancy creation. Each of these components obtains by aggregation of the single-spell partial deadweight losses $O_t^e(\mathbf{y})$, $O_t^\phi(\mathbf{y})$ and $O_t^\theta(\mathbf{y})$, see (39). This appendix reports the $O_t^e(\mathbf{y})$-, $O_t^\phi(\mathbf{y})$- and $O_t^\theta(\mathbf{y})$-values by job ladder rung, and also reports the weights ω_0^i and $\omega_1^i(\mathbf{y})$ used in the aggregation.

F.3.1 Calibrated $O_t^e(\mathbf{y})$-, $O_t^\phi(\mathbf{y})$- and $O_t^\theta(\mathbf{y})$-values

Figure F.3 show single-spell partial deadweight losses from distorted job search effort, job ranking, and vacancy creation. Panels (a), (c), and (e) show single-spell partial deadweight losses in the unemployment submarkets, $O_t^e(\mathbf{y}_0)$, $O_t^\phi(\mathbf{y}_0)$ and $O_t^\theta(\mathbf{y}_0)$, by worker-type. Panels (b), (d), and (f) show single-spell partial deadweight losses in the employed submarkets, which are independent of worker-types in our calibrated economy.

Panel (a) renders the deadweight loss contributions from distorted unemployed job search effort by worker-type and reveals that the higher-type workers generate higher (single-spell partial) deadweight losses. Panel (b) shows deadweight losses from distorted on-the-job job search effort by productivity- and amenity-attributes. Deadweight loss contributions are larger in submarkets where search effort is most distorted and where search effort has the largest impact on the tax base. As is evident in panel (b), these submarkets are predominantly high productivity-, low amenity-attribute submarkets (the north-west corner, where workers search "too much", cf. Figure 6) and, to a lesser extent, low productivity, medium amenity-attribute submarkets (bottom row, middle columns, where workers search "too little", cf. Figure 6).

Panel (c) reflects that, by construction, income taxation does not distort the ranking of jobs among unemployed workers in our calibration. Among employed workers, it is evident from panel (d) that (single-spell partial) deadweight losses arise in submarkets with high productivity-attributes and low amenity attributes (that workers rank "too low", cf. Figure 6), and also among low productivity-, medium amenity-attributes (that workers rank "too high", cf. Figure 6).

Panel (e) shows that distorted vacancy creation in the unemployment submarkets generate deadweight losses that are increasing in worker-type. Panel (f) shows that, in the employed submarkets, distorted vacancy creation gives deadweight losses that are larger in markets with low amenity-attributes (where vacancy creation is "too high", cf. Figure 6), with particularly large deadweight losses in submarkets that also has high productivity-attributes.

F.3.2 Calibrated aggregation weights

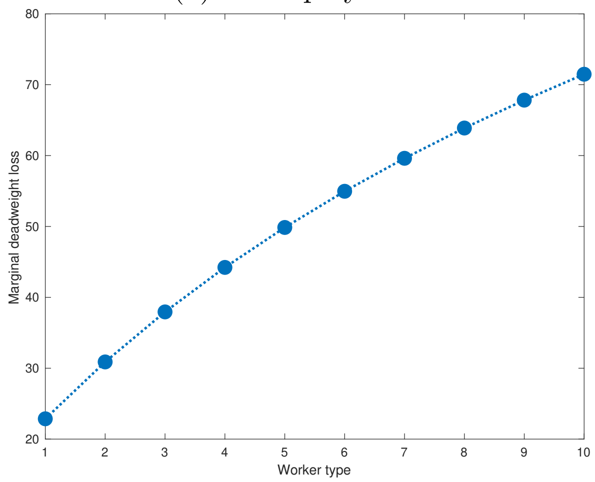
The extent to which the single-spell partial deadweight losses contribute to the overall marginal deadweight loss also depends on the weights ω_0^i and $\omega_1^i(\mathbf{y})$ used in the aggregation, see (39).

Figure F.4 report these weights. Panel (a) shows ω_0^i, the weight put on the unemployed submarket, by worker-type. The weight is increasing in worker-type. Panels (b) and (c) shows the weights put on the employed submarkets for worker-type 1 (the least productive worker-type) and for worker-type 10 (the most productive worker-type), respectively; that is, panels (b) and (c) shows $\omega_1^1(\mathbf{y})$ and $\omega_1^{10}(\mathbf{y})$ by $\mathbf{y}$. These weights are also increasing in worker-type.

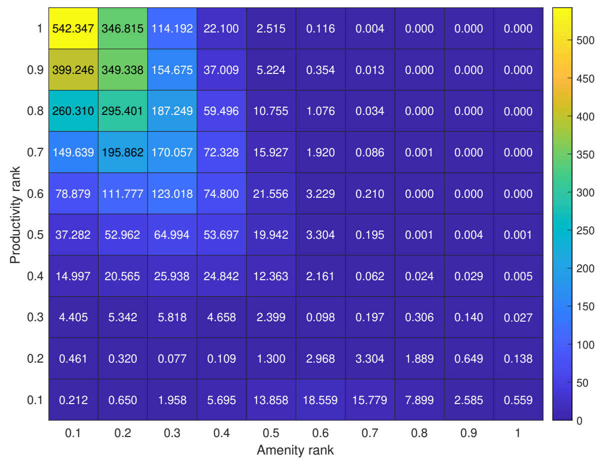
Figure F.3: $O_t^e(\mathbf{y})$, $O_t^\phi(\mathbf{y})$ and $O_t^\theta(\mathbf{y})$ by Worker-Type and Productivity- and Amenity-Attribute

Job search effort, $O_t^e(\mathbf{y})$

(a) Unemployment

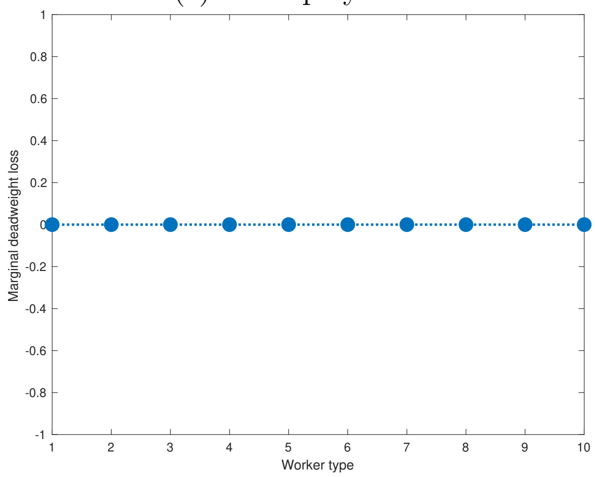


(b) Employment

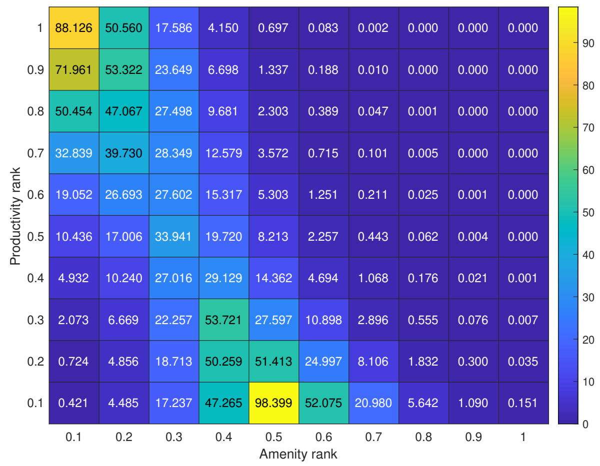


Job ranking, $O_t^\phi(\mathbf{y})$

(c) Unemployment

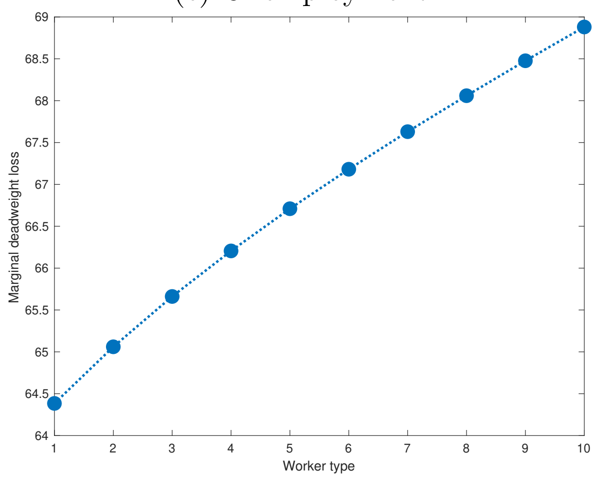


(d) Employment



Vacancy creation, $O_t^\theta(\mathbf{y})$

(e) Unemployment



(f) Employment

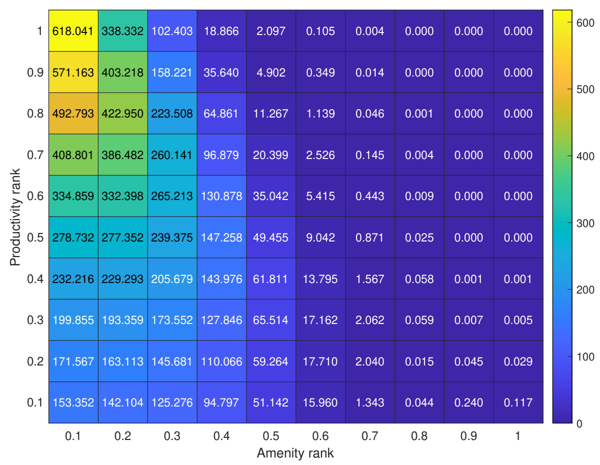
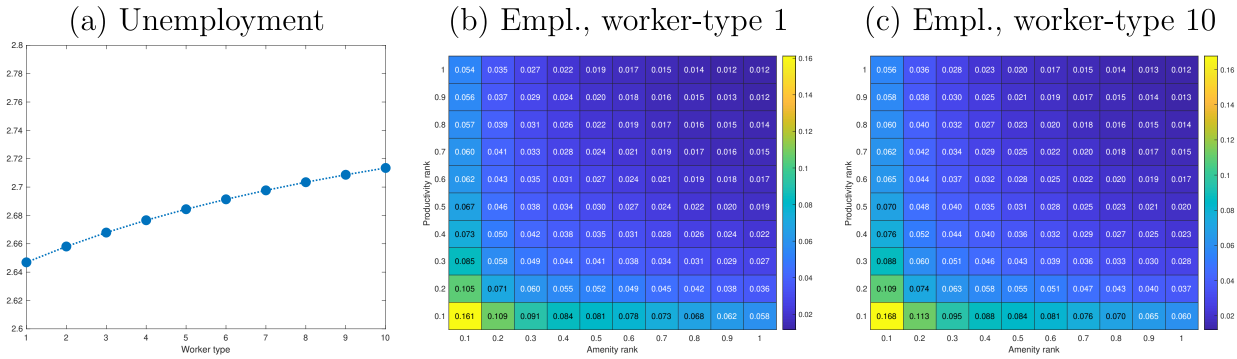


Figure F.4: ω_0^i and $\omega_1^i(\mathbf{y})$ by Worker-Type and Productivity- and Amenity-Attribute



Notes: Panel (a) shows ω_0^i and panel (b) and (c) shows $\omega_1^i(\mathbf{y})$, see (39). Worker-type 1 is the least productive and worker-type 10 is the most productive worker-type.