Economics
Working Papers

2018-6

Income Taxation and the Equilibrium Allocation of Labor

Jesper Bagger, Mads Hejlesen, Kazuhiko Sumiya and Rune Vejlin

Abstract:
We study the impact of labor income taxation on workers' job search behavior and the implications it has for the equilibrium allocation of heterogenous workers across heterogenous firms. The analysis is conducted within a complete markets equilibrium on-the-job search model with two-sided heterogeneity, endogenous job search effort and hiring intensity, equilibrium wage formation, and firm entry and exit. In a nutshell, by appropriating part of the gain from finding a better paid job, income taxation reduces the return to job search effort, and distorts workers' job search effort, which, in turn, distorts the equilibrium allocation of labor. The model is estimated on Danish matched employer-employee data, and is used to evaluate a series of tax reforms in Denmark in the 1990s and 2000s, to provide new insights into the elasticity of taxable labor income, and to identify a Pareto optimal income tax reform.

AARHUS BSS
DEPARTMENT OF ECONOMICS AND BUSINESS ECONOMICS
AARHUS UNIVERSITY

Income Taxation and the Equilibrium Allocation of Labor*

Jesper Bagger†
Royal Holloway and the Dale T. Mortensen Centre

Mads Hejlesen‡
Aarhus University

Kazuhiko Sumiya§
Royal Holloway

Rune Vejlin¶
Aarhus University

August 17, 2018

*We thank Torben M. Andersen, Pierre Cahuc, Rasmus Lentz, Jean-Babtiste Micau, Rob Shimer, seminar participants at Aarhus University, Royal Holloway, CREST, Hitotsubashi University, University of Tokyo, the National Graduate Institute for Policy Studies (Tokyo), ISER/Osaka University, and participants at the 2017 Markets with Search Frictions conference (Aarhus) and the Yale "Macro Breakfast" PhD workshop for helpful comments and suggestions. We gratefully acknowledge support from The Danish Council for Independent Research — Social Sciences, grant DFF 6109-00037

†Department of Economics, Royal Holloway, University of London, Egham, Surrey, TW20 0EX, United Kingdom; E-mail: jesper.bagger@rhul.ac.uk

‡Department of Economics and Business, Aarhus University, Fuglesangs Allé 4, 8210 Aarhus V, Denmark; E-mail: mhejlesen@econ.au.dk

§Department of Economics, Royal Holloway, University of London, Egham, Surrey TW20 0EX, United Kingdom; E-mail: kazuhiko.sumiya.2014@live.rhul.ac.uk

¶Department of Economics and Business, Aarhus University, Fuglesangs Allé 4, 8210 Aarhus V, Denmark; E-mail: rvejlin@econ.au.dk

Abstract

We study the impact of labor income taxation on workers' job search behavior and the implications it has for the equilibrium allocation of heterogenous workers across heterogenous firms. The analysis is conducted within a complete markets equilibrium on-the-job search model with two-sided heterogeneity, endogenous job search effort and hiring intensity, equilibrium wage formation, and firm entry and exit. In a nutshell, by appropriating part of the gain from finding a better paid job, income taxation reduces the return to job search effort, and distorts workers' job search effort, which, in turn, distorts the equilibrium allocation of labor. The model is estimated on Danish matched employer-employee data, and is used to evaluate a series of tax reforms in Denmark in the 1990s and 2000s, to provide new insights into the elasticity of taxable labor income, and to identify a Pareto optimal income tax reform.

Keywords: Labor reallocation, Income taxation, Tax reforms, Worker heterogeneity, Firm heterogeneity, Matched employer-employee data

JEL codes: H20, J30, J64, J63

1 Introduction

Labor income taxation is a key vehicle by which governments redistribute income, alleviate risk and finance public goods and services. Indeed, developed economies raise 35%-50% of national income in taxes, with labor income taxation being the primary source of revenue, accounting for roughly 75% of the total tax burden (Piketty and Saez, 2013). However, at the same time, taxation distorts incentives, leaves gains from trade unexploited, and hampers economic efficiency and growth. The design of optimal income tax schedules that achieve redistributive goals with minimum distortionary effects requires detailed knowledge of how economic agents' respond to taxation. As a result, a large and diverse empirical literature studying behavioral responses to income taxation has emerged.[1] We contribute to this literature by studying the impact of labor income taxation on workers' job search behavior and the implications it has for the equilibrium allocation of heterogenous workers across heterogenous firms.

Our analysis is based on a rich equilibrium on-the-job search model with two-sided heterogeneity, endogenous job search effort on the worker-side, as in Burdett (1978), Christensen, Lentz, Mortensen, Neumann, and Werwatz (2005), and Bagger and Lentz (2015), endogenous hiring intensity on the firm-side, wage determination as in Burdett and Mortensen (1998) and Bontemps, Robin, and van den Berg (2000), and firm entry and exit as in Melitz (2003). Curvature in the utility of consumption implies that workers' optimal search effort choices feature both substitution and income effects, and introduces a desire for consumption smoothing in the face of labor income risk stemming from job finding and job destruction events. We assume the presence of a complete system of insurance markets in which workers may trade securities to alleviate this consumption risk, and solve for the complete markets equilibrium using a large household formulation of the model in the mould of Merz (1995). In a nutshell, the model stipulates that income taxation reduces the return to job search effort, thus distorting workers' job search effort, and therefore the equilibrium allocation of heterogenous workers across heterogenous firms.

The empirical analysis makes use of a comprehensive population-wide Danish matched-employer employee panel covering 1990-2005. The data includes detailed information on workers individual tax liabilities, which allow us to obtain precise estimates of the changing Danish tax system during the observation period. We structurally estimate the equilibrium model by way of Indirect Inference using data pertaining to 1999-2003. The estimated model is used for three exercises that illustrate and quantify the impact of labor income taxation on the equilibrium allocation of labor. First, we evaluate how three Danish income tax reforms during 1990-2005 affected the steady state unemployment rate and taxable labor income. Second, we provide new estimates of a mobility-based long-run elasticity of taxable labor income, including a novel structural decomposition of the elasticity of taxable labor income. Third, we characterize a

[1] See e.g. the Mirrlees Review (Adam, Besley, Blundell, Bond, Chote, Johnson, Myles, and Poterba, 2011, Mirrlees, Adam, Besley, Blundell, Bond, Chote, Johnson, Myles, and Poterba, 2011) for a recent and comprehensive collection of studies of issues related to taxation, including labor income taxation.

Pareto optimal income tax reform, which increases government revenue while leaving all workers (weakly) better off vis-a-vis the baseline 1999-2003 income tax regime in terms of steady state expected utility.

We find that the 1994, 1999, and 2004 Danish income tax reforms improved the equilibrium allocation of labor, reducing the steady state unemployment rate by about 6 percent, and increasing steady state labor income by about 2 percent, with workers of low ability gaining the most in terms of lower unemployment rates and higher labor incomes. Currently, our estimation procedure and subsequent quantitative analysis does not balance the government budget, and overall, the tax reforms during 1990-2005, lead to a 2 percent drop in government revenue. Accounting for equilibrium adjustments is quantitatively important when evaluating the three income tax reforms.

The estimated model implies that the full equilibrium mobility-based elasticity of taxable labor income is 0.12. The corresponding compensated elasticity of taxable labor income, where any income effects have been eliminated, is larger, at 0.17. Moreover, we find that workers of lower ability have higher taxable labor income elasticities. A structural decomposition reveal that 46 percent of the measured elasticity of taxable income come from extensive margin allocative adjustments (i.e. changes in the steady state unemployment rate), 12 percent come from intensive margin allocative adjustments (reflecting changes in the steady state distribution of employed workers across firms), and 42 percent come from firms' wage policy adjustments. When we apply the decomposition across the ability distribution, we see that extensive margin adjustments drives the bulk of the measured taxable labor income elasticity at the bottom of the ability distribution, whereas intensive margin adjustments and adjustment to wage policies drive the taxable labor income elasticity estimates towards the top of the ability distribution.

Finally, we identify a Pareto optimal income tax reform, see Werning (2007), Blundell and Shephard (2012) and Hosseini and Shourideh (2017), which increases government revenue by 0.68 percent while leaving no workers worse off compared to the baseline 1999-2003 income tax schedule. The Pareto optimal income tax schedule reduces the marginal tax rate at low wage levels, and increases it slightly at all higher wage levels. At the same time, the average tax rate is increased at low wages, and reduced at higher wage levels. Interestingly, the Pareto optimal tax schedule shares some features of the actual tax reforms implemented in Denmark during the 1990-2005 period.

There are only a few other papers providing empirical evidence on the effect of income taxation on job search effort, labor mobility and labor allocation. Gentry and Hubbard (2004) presents a reduced form regression analysis of the impact of income taxation on workers' labor market mobility. Using time- and state-variation in income taxation in the US as a source of identification, Gentry and Hubbard (2004) find that a 5 percentage point reduction in the marginal tax rate increases the probability that a worker (in their study, a head of household) is in better job within a year by 0.79 percentage points, and that a reduction in income tax progressivity increases the probability that a worker moves to better job. In a paper closely related to ours, Kreiner, Munch, and Whitta-Jacobsen (2015) estimate a partial equilibrium on-the-job search model on Danish data in order to compute a mobility-based elasticity of taxable labor income. Kreiner, Munch, and Whitta-Jacobsen (2015) finds a taxable labor income elasticity of 0.30, but their analysis does not account for income effects in job search behavior, does not explicitly account for both worker and firm heterogeneity, and does not allow equilibrium adjustments to changes in the income tax system. Shephard (2017) develops an equilibrium search model where firm's differ in productivity, and workers differ in the value of non-market time. The model, which also features both part- and full-time jobs, is estimated using UK data, including a detailed accounting for income taxation, and is then used to conduct an equilibrium evaluation of the labor supply responses to a reform of the Working Families Tax Credit, a UK in-work benefit scheme. Shephard (2017) does not have endogenous search on the worker side and focus instead on the labor supply responses at the part-time and full-time margin, and does not use matched employer-employee data in the estimation. Breda, Haywood, and Wang (2017) develops an equilibrium search model to analyze labor market effects of taxation and minimum wage policies in France, but do not endogenize on-the-job search intensity, and thus primarily focus on workers' incentive to leave unemployment.

The paper proceeds as follows. Section 2 presents the model and characterizes the equilibrium, section 3 describes the data we use, section 4 provides a detailed description of the Danish income tax system during the relevant 1990-2005 period and also describes how we estimate the tax function used in the estimation. Section 5 details the estimation procedure, including a discussion of the estimated model's fit to data, and a presentation of the estimated structural parameters, section 6 contains the equilibrium evaluation of the 1994, 1999, and 2004 Danish income tax reforms, section 7 contains our analysis of the elasticity of taxable income, and section 8 presents a Pareto optimal income tax reform based on the estimated model. Finally, section 9 concludes, and a number of appendicies contain technical details and supplementary results.

2 The model

2.1 The environment

The model is set in continuous time and the future is discounted at rate ρ. We assume throughout that the labor market is in steady state.

Two-sided heterogeneity. One side of the labor market is populated by a unit measure of infinitely lived individuals of heterogenous scalar ability $a \in [0,1]$. Let $H(\cdot)$ be the CDF of ability in the worker population. We take $H(\cdot)$ to be continuously differentiable with PDF $h(\cdot)$. On the other side of the market, there is a measure M_0 firms of heterogenous productivity $p \in [p_0, 1]$. Productivity p is distributed with CDF $\Gamma_0(\cdot)$ across firms. M_0, p_0 and $\Gamma_0(\cdot)$ are endogenous objects.

Match output. A match between an ability-a worker and a productivity-p firm yields output $y(a,p)$. The match production function $y(a,p)$ is assumed twice continuously differentiable with $y'_a(a,p) > 0$ and $y'_p(a,p) > 0$ for all (a,p)-combinations. All other derivatives of the match output function are left unrestricted. The total output of a firm is the sum of the output from the matches it participates in. That is, the production function at the firm-level exhibits constant returns to scale in matches.

Match formation and dissolution. We assume labor markets are segregated by worker-ability a. Search frictions require workers and firms to exert effort to form productive matches, and workers may search both off- and on-the-job. A type-a worker who exerts job search effort s receives job offers from hiring firms at rate $\lambda(a)s$, where $\lambda(a)$ is the job finding rate per unit of search effort for type-a workers, which is an equilibrium object that depends on the tightness of the market for type-a labor.

A job offer is a draw of a gross wage w from the ability-conditional wage offer distribution $F(w|a)$ with support $[\underline{w}(a), \overline{w}(a)]$. When a worker accepts a job offer, a match is formed, and the worker starts working for the new employer immediately. Matches are dissolved in one of three ways. First, workers are laid off at rate $\delta(a)$ upon which they transition into unemployment. Second, workers quit for higher paying job found through on-the-job search. This event occurs at rate $\lambda(a)s\overline{F}(w|a)$, where s is the worker's endogenous search effort, and where $\overline{F}(w|a) \equiv 1 - F(w|a)$. Third, workers are exogenously reallocated to new matches at rate $\lambda(a)\mu$, with a wage drawn from $F(w|a)$. The worker can reject a reallocation and instead transition to unemployment. Descriptively, the reallocation shocks generates transitions from high to low paying jobs, events that are empirically prevalent, but that the model would otherwise be unable to generate.[2] To simplify the model solution, unemployed workers are also subject to reallocation shocks.

Taxation and the government's budget. A worker's net wage resulting from a gross wage of w is $w - T(w)$, where $T(\cdot)$ is the income tax function. Taxes are paid continuously, and we assume that $T(w) < w$, that $T(\cdot)$ is twice differentiable, and that $0 \leq T'(w) < 1$. Firm profit is taxed at rate τ. An unmatched worker receives gross unemployment benefit b, and the net flow income of an unemployed worker is $b - T(b)$. If the government's steady state revenue from levying income and profit taxes is R, the government budget is the identity

$$B = R - n^0 b - \bar{b} \equiv 0, \tag{1}$$

where, anticipating some notation to be introduced below, n^0 is the aggregate steady state unemployment rate, and $\bar{b}$ is a per-capita lump-sum transfer that balance the government

[2] In terms of interpretation, transitions from high to low paying jobs may reflect advanced layoff notices, strong duration dependence in the unemployment-to-job hazard, or simply considerations that fall outside the scope of the model, e.g. amenities.

budget.[3]

Worker preferences. Workers have preferences over consumption of a composite private good c and job search effort s, an economic bad that yields disutility. The preferences are represented by the additively separable felicity function $\psi : [0, \infty) \times [\underline{s}, \infty) \mapsto \mathbb{R}$, with

$$\psi(c, s) = u(c) - \zeta(s - \underline{s}), \tag{2}$$

where $u : [0, \infty) \mapsto \mathbb{R}$ is a continuously differentiable function with $u'(\cdot) > 0$ and $u''(\cdot) \leq 0$, and $\zeta : [0, \infty) \mapsto [0, \infty)$ is a continuously differentiable function with $\zeta'(\cdot) > 0$ and $\zeta''(\cdot) > 0$. This parameterization bounds search effort from below by $\underline{s} > 0$, reflecting that low levels of search effort, up until $s = \underline{s}$, involves no disutility. We refer to $\underline{s}$ as "free search". Free search turns out not to be quantitatively important, but rules out degenerate equilibrium wage offer distributions (Diamond, 1971). As we shall see, the convexity of $\zeta(\cdot)$ determines the substitution effect in the job search response to income tax reforms. The income effect is pinned down by the concavity of $u(\cdot)$. Insofar as the job search literature has endogenized search effort, it typically imposes quasi-linear preferences that rules out income effects in job search behavior.[4] With quasi-linear preferences, a reduction in the marginal income tax rate invariably result in increased worker search effort. In the more general case of (2), search effort may be reduced if the income effect dominates the substitution effect.

Large households and the complete markets allocation. Curvature in the utility of consumption, $u(\cdot)$, also introduces a desire for consumption smoothing in the face of labor income risk stemming from job finding and job destruction events. We deal with this issue within the complete markets paradigm, and follow Merz (1995) in characterizing the resulting allocation (of consumption and search effort) by introducing a set of large households in the economy. Specifically, each worker in the economy belong to large household that is able to dictate the job search effort of its members. Members of a household pool their income to provide perfect consumption insurance against labor income fluctuations. A household consists of a unit measure of workers of identical ability. We refer to a household of ability-a workers as a type-a household. There are $h(a)$ type-a households, coinciding with the measure of ability-a workers, and the economy thus contains a unit measure of households, $\int_0^1 h(a)da = 1$.

Even though our analysis is strictly positive—measuring and analyzing the consequences of the actual tax schedule in place and side-stepping any normative consideration regarding the shape of this tax schedule—we note here that the complete markets assumption eliminates the

[3] As will become clear as we present the model, imposing a balanced budget is not a trivial exercise and it involves some numerical complications that we have yet to solve in a satisfactory way. For this reason, our current estimation procedure and analysis is done with $\bar{b} = 0$.

[4] Indeed, existing empirical work on the effect of income taxation on job search and job mobility either takes no account of income effects (see Kreiner, Munch, and Whitta-Jacobsen, 2015), or does not separate income and substitution effects (Gentry and Hubbard, 2004). This is in sharp contrast to the empirical labor supply literature, see e.g. Heckman (1993) and Blundell and MaCurdy (1999).

need for government-provided social insurance through progressive taxation. However, as our model features ex-ante heterogeneous workers, a progressive income tax schedule may nonetheless be warranted out of concerns for redistribution. Heathcote, Storesletten, and Violante (2017) provide a thorough analysis of the factors that shape the optimal degree of income tax progressivity. Michau (2017) analyze optimal provision of social insurance in an on-the-job search model of ex-ante homogenous workers.

Wage determination. The ability-conditional wage offer distribution $F(w|a)$ is determined in a wage posting equilibrium as in Burdett and Mortensen (1998) and Bontemps, Robin, and van den Berg (2000). The wage setting game restricts admissible wage policies in three dimensions. First, firms make take-it-or-leave-it offers and credibly commit to future wage payments. Second, the wage must remain constant throughout the duration of the match. Third, contracts are anonymous, implying that a firm must offer the same contract to all workers of the same ability. A firm's wage policy maximizes its steady state profit by trading off profit per worker and worker turnover, and is a best response to all other firms' wage and recruitment policies and workers' optimal search effort choices. The equilibrium outcome is a set of non-degenerate ability-specific wage offer distributions, where more productive firms offer higher wages.

Alternative wage setting arrangements are available. Coles (2001) characterize a wage posting equilibrium where firms may renege on promised wage payments. Stevens (2004), and Burdett and Coles (2003) and Burdett and Coles (2010) relax the assumption that the offered wage remain constant throughout the match and characterize the set of (privately) optimal wage-tenure contracts. The contracts derived in Stevens (2004) has little empirical content. Burdett and Coles (2003) and Burdett and Coles (2010) consider risk-averse workers who inhabit an incomplete markets economy where there is value to firms smoothing workers consumption paths through the wage contract. While the resulting wage contracts are empirically plausible, with complete markets, there is no need for firms to provide such insurance. Postel-Vinay and Robin (2002) and Cahuc, Postel-Vinay, and Robin (2006) relax the assumption of take-it-or-leave-it offers through the introduction of sequential auctions and bargaining in the presence of on-the-job search.[5] Lentz (2010) and Bagger and Lentz (2015) extend the Cahuc, Postel-Vinay, and Robin (2006) framework to allow for endogenous job search effort. However, with income and profit taxation, wages are no longer pure transfers in the bargaining game, rendering the Cahuc, Postel-Vinay, and Robin (2006) framework unattractive for our purposes.[6] Finally, Lentz (2015)

[5] From a theoretical point of view, Postel-Vinay and Robin (2004) consider whether firms should match outside offers and reach the (tentative) conclusion that more productive firms may benefit from an offer-matching strategy, whereas firms with lower productivity may be better off committing not to match outside offers. From an empirical point of view, using survey data, Hall and Krueger (2012) find that about 1/3 of the workers in their survey have wages set through take-it-or-leave-it offers, whereas another 1/3 were able to bargain.

[6] Breda, Haywood, and Wang (2017) uses a sequential bargaining model in the mould of Cahuc, Postel-Vinay, and Robin (2006) to analyze the effect of payroll taxes on labor allocation, but does not provide a solution to the wage bargaining game, instead imposing a reduced form sharing rule.

studies equilibrium wage formation in an incomplete markets economy where on-the-job search effort is subject to moral hazard.

2.2 Household behavior

Consider a type-a household. Let $n_t^0(a)$ be the measure of unemployed household members at time t, and let $n_t^1(w|a)$ be the measure of household members employed with a gross wage w at time t. Define

$$N_t(a) \equiv n_t^0(a) + \int_{\phi(a)}^{\overline{w}(a)} n_t^1(w|a)dw \tag{3}$$

as the time-t measure of the household population.

The distribution of workers across unemployment and employment, and within employment, across wage levels depend on household members' labor market behavior. The household is able to dictate, and enforce, the search effort and acceptance decisions of its members. Let $s_t^0(a)$ be the dictated time-t search intensity for an unemployed type-a worker, and let $s_t^1(w|a)$ be the dictated time-t search intensity for a type-a worker with gross wage w. In regard to the dictated acceptance decision of job offers, we confirm below that the household will employ a reservation wage strategy. Let $\phi(a)$ be the reservation wage dictated to unemployed workers in a type-a household. The reservation wage of members employed with wage w is, obviously, w.

Members of the household pool their net-of-tax labor income and perfectly insure each other against consumption fluctuations due to individual job finding and job destruction shocks. Hence, time-t consumption of a type-a household member is independent of current individual labor market status and labor income, and is given by

$$c_t(a) = \frac{1}{N_t(a)}\left([b - T(b)]n_t^0(a) + \int_{\phi(a)}^{\overline{w}(a)} [w - T(w)]n_t^1(w|a)dw\right) + \bar{b}, \tag{4}$$

the average type-a household net-of-tax income at time t. Recall that $\bar{b}$ is a government lump-sum transfer that balances the government budget.

Turning to the labor market behavior, a type-a household set unemployed search effort profile $s_t^0(a)$, an employed search effort profile $s_t^1(w|a)$ for every $w \in [\phi(a), \overline{w}(a)]$, as well as a reservation wage $\phi(a)$, in order to maximize the present value household lifetime utility, $\Psi(a)$, given as

$$\Psi(a) = \int_0^\infty e^{-\rho t}\left[\psi(c_t(a), s_t^0(a))n_t^0(a) + \int_{\phi(a)}^{\overline{w}(a)} \psi(c_t(a), s_t^1(w|a))n_t^1(w|a)dw\right]dt, \tag{5}$$

where $\psi(c, s)$ is the felicity function given by (2), and $c_t(a)$ is the household consumption level given by (4). The optimization is subject to the law of motions for $n_t^0(a)$ and $n_t^1(w|a)$. The law of motion for $n_t^0(a)$ is given by

$$\dot{n}_t^0(a) = \delta(a)N_t(a) - [\delta(a) + \lambda(a)\mu + \lambda(a)s_t^0(a)]n_t^0(a), \tag{6}$$

The "dot"-notation indicates derivatives with respect to time t. The law of motion for $n_t^1(w|a)$ is slightly more involved and is given by

$$\dot{n}_t^1(w|a) = \left[\lambda(a)\mu N_t(a) + \lambda(a)s_t^0(a)n_t^0(a) + \int_{\phi(a)}^{w} \lambda(a)s_t^1(x|a)n_t^1(x|a)dx\right] f(w|a) - \left[\delta(a) + \lambda(a)\mu + \lambda(a)s_t^1(w|a)\overline{F}(w|a)\right] n_t^1(w|a), \quad (7)$$

for every $w \in [\phi(a), \overline{w}(a)]$. The derivation of both (6) and (7) are based on gross worker flow accounting and are detailed in Appendix A.1.

Conditional on $\phi(a) < \overline{w}(a)$, the household problem is an optimal control problem with the controls $s_t^0(a)$ and $s_t^1(w|a)$ for every $w \in [\phi(a), \overline{w}(a)]$, and states $n_t^0(a)$ and $n_t^1(w|a)$ for every $w \in [\phi(a), \overline{w}(a)]$. Define $\xi_t^0(a)$ as the multiplier on the law of motion for $n_t^0(a)$ in the household's problem, and let $\xi_t^1(w|a)$ be the multiplier on the law of motion for $n_t^1(w|a)$. $\xi_t^0(a)$ and $\xi_t^1(w|a)$ are referred to as the costate variables, and we denominate them in present ($t = 0$) values. The costates can be interpreted as shadow values, and are central to the analysis to come. Indeed, $\xi_t^0(a)$ is the increase in the present value household lifetime utility $\Psi(a)$, see (5), from adding an additional unemployed worker to the household at time t. Likewise, $\xi_t^1(w|a)$ is the increase in $\Psi(a)$ from adding an extra worker with wage w to the household at time t.

Our analysis focuses on an economy in steady state, where controls, states and the *current value* costates, given by $e^{\rho t}\xi_t^0(a)$ and $e^{\rho t}\xi_t^1(w|a)$, are time-invariant, and the *present value* costates, $\xi_t^0(a)$ and $\xi_t^1(w|a)$, varies with t only because of discounting. That is, steady state behavior is captured by $\xi_0^0(a)$ and $\xi_0^1(w|a)$ with the following set of steady state identities applying: $\xi_0^0(a) = e^{\rho t}\xi_t^0(a)$ and $\xi_0^1(w|a) = e^{\rho t}\xi_t^1(w|a)$ for every $w \in [\phi(a), \overline{w}(a)]$.[7] Steady state values of $s_t^0(a)$ and $s_t^1(w|a)$ are denoted by $s^0(a)$ and $s^1(w|a)$, respectively.

Suppose an unemployed household member has received an offer of a job paying a wage w. In order to maximize $\Psi(a)$, the household dictates a rejection if the value of the worker remaining unemployed, $\xi_0^0(a)$, exceeds the value of having the worker employed in a job paying w, $\xi_0^1(w|a)$. The household dictates acceptance if the opposite inequality between $\xi_0^0(a)$ and $\xi_0^1(w|a)$ holds. We confirm below that $\xi_0^1(w|a)$ is strictly increasing in w. Hence, the household employs a reservation wage strategy, and the current reservation wage $\phi(a)$ solves

$$\xi_0^1(\phi(a)|a) = \xi_0^0(a). \quad (8)$$

The household's optimal control problem can be solved by constructing the associated Hamiltonian and applying a version of the Maximum Principle (Acemoglu, 2009, section 7.5, p. 253). We state just the main results here, with details on the derivations relegated to Appendix A.3. Optimal steady state search effort for employed workers earning a wage $w \in [\phi(a), \overline{w}(a)]$ satisfies

$$\zeta'(s^1(w|a) - \underline{s}) = \lambda(a)\int_{w}^{\overline{w}(a)} [\xi_0^1(x|a) - \xi_0^1(w|a)]dF(x|a) = \lambda(a)\int_{w}^{\overline{w}(a)} \frac{\partial}{\partial x}\xi_0^1(x|a)\overline{F}(x|a)dx, \quad (9)$$

[7] $\xi_0^0(a)$ and $\xi_0^1(w|a)$ are the added value to $\Psi(a)$ from immediately admitting an additional unemployed worker, respectively employed worker with wage w, to the household when the economy is in steady state.

where the last equality follows from integrating by parts. The interpretation of (9) is straightforward. Suppose the household perturb its steady state search policy by dictating marginally higher search intensity at wage-rung w, specifically from $s^1(w|a)$ to $s^1(w|a) + \Delta$. This pertubation involves an additional disutility-flow $\zeta'(s^1(w|a) - \underline{s})\Delta$ per rung-w worker, see (2). However, it also increases the rate at which rung-w workers obtain job offers by $\lambda(a)\Delta$. The additional offers are distributed according to $F(\cdot|a)$, but only a share $\overline{F}(w|a)$ are acceptable. Reallocating a worker from rung w to rung $x > w$ yields a household utility gain of $\xi_0^1(x|a) - \xi_0^1(w|a)$. Hence, the expected per-worker marginal value gain of the search policy perturbation is $\lambda(a)\Delta \int_w^{\overline{w}(a)} [\xi_0^1(x|a) - \xi_0^1(w|a)]dF(x|a)$. Condition (9) therefore balances the added marginal disutility from search with the expected marginal gains from search. An analogous condition characterizes optimal unemployed search.[8]

A recursive expression for the steady state value to the household of a worker employed at wage w, $\xi_0^1(w|a)$, can be derived as

$$\rho\xi_0^1(w|a) = \psi(c(a), s^1(w|a)) + u'(c(a))[w - T(w) - c(a)] + \delta(a)[\xi_0^0(a) - \xi_0^1(w|a)] + \lambda(a)\mu \int_{\phi(a)}^{\overline{w}(a)} \frac{\partial}{\partial x}\xi_0^1(x|a)\overline{F}(x|a)dx + \lambda(a)s^1(w|a) \int_w^{\overline{w}(a)} \frac{\partial}{\partial x}\xi_0^1(x|a)\overline{F}(x|a)dx, \quad (10)$$

where $\psi(\cdot,\cdot)$ is the felicity function given by (2), and where $c(a)$ is the steady state version of household consumption, see (4). A worker employed with wage w is an asset to the household. The left-hand side of (10) is the permanent utility flow from this asset. This flow comprises the household payoff flow, and the expected returns to job destruction, job reallocation, and job finding. The marginal wage-rung w worker increases the household utility flow by worker-felicity $\psi(c(a), s^1(w|a))$, which leaves net-of-tax revenue in the amount of $w - T(w) - c(a)$ to be distributed within the household. The household values this additional revenue at the marginal utility of consumption, $u'(c(a))$. Job destruction, job reallocation, and job finding events are associated with capital gains from the worker transitioning into unemployment, being reallocated to a randomly drawn new job, or being matched with a new potential employer through on-the-job search.

In Appendix A.3 we derive a similar expression for the steady state value of an unemployed worker, $\xi_0^0(a)$, which reads

$$\rho\xi_0^0(a) = u(c(a)) - \zeta(s^0(a)) + u'(c(a))[b - T(b) - c(a)] + \lambda(a)[\mu + s^0(a)] \int_{\phi(a)}^{\overline{w}(a)} \frac{\partial}{\partial x}\xi_0^1(x|a)\overline{F}(x|a)dx. \quad (11)$$

Substituting (10) and (11) into the reservation wage definition (8) shows that

$$\phi(a) = b. \quad (12)$$

[8]Optimal steady state unemployed search effort $s^0(a)$ solves $\zeta'(s^0(a) - \underline{s}) = \lambda(a) \int_{\phi(a)}^{\overline{w}(a)} [\xi_0^1(x|a) - \xi_0^0(a)]dF(x|a)$.

This result stems from the assumption that unemployed and employed job search is equally efficient, and involves equal disutility. In this case, there are no opportunity cost of taking up a job. We impose $\phi(a) = b$ in the remainder of the paper. Given (12), it is evident that optimal off-the-job search equals optimal on-the-job search in a firm offering $w = b$, that is, $s^0(a) = s^1(b|a)$.[9] This result allow us to simplify notation by dropping the superscript "0" and "1" that has distinguished unemployed and employed search up until now. In the remainder of the paper we let $s(w|a)$ denote optimal search effort at wage w, with $s(b|a)$ being optimal unemployed search.

Taking the derivative of (10) with respect to w, applying the Envelope Theorem, and substituting the resulting expression for $\partial\xi_0^1(w|a)/\partial w$ into (9) yields the following expression for the complete markets optimal job search effort as a function of the current wage w,

$$\zeta'(s(w|a) - \underline{s}) = \lambda(a) \int_w^{\overline{w}(a)} \frac{u'(c(a))[1 - T'(x)]\overline{F}(x|a)}{\rho + \delta(a) + \lambda(a)\mu + \lambda(a)s(x|a)\overline{F}(x|a)} dx. \tag{13}$$

The optimal search effort balances the marginal cost with the marginal gains from on-the-job search. In general, workers of different ability differ in their chose job search effort, due to differences in the return to job search stemming from differences in the distribution of job offers $\overline{F}(x|a)$ across workers of different abilities. Lemma 1 summarizes some useful properties of the optimal search effort.

Lemma 1. *Optimal search effort $s(w|a)$ has the following properties (i) $s(w|a) \geq \underline{s}$, (ii) $s(\overline{w}(a)|a) = \underline{s}$, and (iii) $\partial s(w|a)/\partial w < 0$ for every $w \in [b, \overline{w}(a))$.*

Proof. See Appendix A.4. □

Lemma 1 states that job search effort is bounded from below by $\underline{s} > 0$. This is a useful feature of the model as it rules out degenerate equilibrium wage offer distributions, see Diamond (1971), Burdett and Judd (1983), and Burdett and Mortensen (1998). The return to job search effort diminishes at higher wages because there are only limited possibilities to climb further up the wage ladder. Indeed, at the highest paying job, the returns to on-the-job search is zero, and no job search effort is forthcoming. The model implies that unemployed workers exert more job search effort than employed workers, despite facing the same disutility of effort and the same search efficiency. The impact of income taxation on job search is discussed further below when we consider partial equilibrium substitution and income effects in the provision of job search effort. We note, however, that (13) implies that search effort is distorted by income taxation, and it is the average tax (i.e. the integrated marginal tax rate) that matters for distortions.[10] Furthermore, income taxation may have different quantitative implications for the job search behavior of workers of different abilities. These effects of income taxation will all be born out in the empirical analysis to come.

[9] For details, consult Appendix A.3.

[10] Integration by parts of (13) results in a differential equation for $s(w|a)$ that does not depend on the marginal tax function $T'(\cdot)$, but only the average tax function $T(\cdot)$.

2.2.1 The complete markets steady state labor allocation

Normalizing the steady state household population $N_t(a) \equiv 1$ allow us to interpret $n^0(a)$ as the steady state unemployment rate among type-a workers, and to define the steady state PDF of employed workers across wage rates as $g(w|a) \equiv n^1(w|a)/[1 - n^0(a)]$, with associated CDF $G(w|a) = \int_b^{\overline{w}} g(w|a)dw$. Together, $n^0(a)$ and $G(w|a)$ characterizes the steady state allocation of type-a labor.

Using the steady state version of (6) with $N_t(a) \equiv 1$ implies that

$$n^0(a) = \frac{\delta(a)}{\delta(a) + \lambda(a)\mu + \lambda(a)s(b|a)}, \tag{14}$$

where $s(b|a)$ is given by (13). Furthermore, the steady state version of (6) with $\phi(a) = b$ and $N_t(a) \equiv 1$ yields the following balanced flow equation defining $G(w|a)$,

$$[\lambda(a)\mu + \lambda(a)s^0(a)]F(w|a)\frac{n^0(a)}{1 - n^0(a)} + \lambda(a)\mu F(w|a)\overline{G}(w|a)$$
$$= \delta(a)G(w|a) + \overline{F}(w|a)\int_b^w \lambda(a)s(x|a)g(x|a)dx. \tag{15}$$

The left-hand side of reflects the inflow into the stock of workers earning a wage w or less, and the right-hand side reflects the outflow. In steady state, in- and outflow balance. We show in Appendix A.2 that (15) implies the following closed-form expression for the allocation of labor,

$$G(w|a) = 1 - \exp\left\langle -\int_b^w \left[\frac{\delta(a) + \lambda(a)\mu}{\delta(a) + \lambda(a)\mu + \lambda(a)s(x|a)\overline{F}(x|a)}\right]\frac{f(x|a)}{\overline{F}(x|a)}dx\right\rangle. \tag{16}$$

Since all type-a households are identical, $G(w|a)$ represents both the allocation of labor within a type-a household, as well as the economy-wide allocation of type-a workers.

2.2.2 Tax reforms and partial equilibrium substitution and income effects

The analysis to follow is based on a comparative statics analysis of steady state partial equilibrium predictions of the impact of income taxation on job search effort. That is, we compare the steady state job search policies under two different tax regimes, with no account of the out-of-steady state adjustment process.

Consider a tax reform that changes the income tax function from $T(\cdot)$ to $\hat{T}(\cdot)$. Let $s(w|a)$ be the pre-reform job search effort of a type-a worker employed with wage w given by (13), and let $\hat{s}(w|a)$ be the post-reform search effort defined by the post-reform version of (13),

$$\zeta'(\hat{s}(w|a)) = \lambda(a)\int_w^{\overline{w}(a)} \frac{u'(\hat{c}(a))[1 - \hat{T}'(x)]\overline{F}(x|a)}{\rho + \delta(a) + \lambda(a)\mu + \lambda(a)\hat{s}(x|a)\overline{F}(x|a)}dx, \tag{17}$$

where $\hat{c}(a)$ is post-reform consumption, computed according to the steady state version of (4) imposing the post-reform tax system and the post-reform labor allocation. Finally, let $\tilde{s}(w|a)$ be the job search intensity of a type-a worker employed with wage w under the new tax regime

Economics Working Papers

2017-07: Ina C. Jäkel Allan Sørensen: Exporter Price Premia?

2017-08: Marianne Simonsen, Lars Skipper and Niels Skipper: Piling Pills? Forward-Looking Behavior and Stockpiling of Prescription Drugs

2017-09: Federico Ciliberto and Ina C. Jäkel: Superstar Exporters: An Empirical Investigation of Strategic Interactions in Danish Export Markets

2017-10: Anna Piil Damm, Britt Østergaard Larsen, Helena Skyt Nielsen and Marianne Simonsen: Lowering the minimum age of criminal responsibility: Consequences for juvenile crime and education

2017-11: Erik Strøjer Madsen: Branding and Performance in the Global Beer Market

2017-12: Yao Amber Li, Valerie Smeets and Frederic Warzynski: Processing Trade, Productivity and Prices: Evidence from a Chinese Production Survey

2017-13: Jesper Bagger, Espen R. Moen and Rune M. Vejlin: Optimal Taxation with On-the-Job Search

2018-01: Eva Rye Johansen, Helena Skyt Nielsen and Mette Verner: Long-term Consequences of Early Parenthood

2018-02: Ritwik Banerjee, Nabanita Datta Gupta and Marie Claire Villeval: Self Confidence Spillovers and Motivated Beliefs

2018-03: Emmanuele Bobbio and Henning Bunzel: The Danish Matched Employer-Employee Data

2018-04: Martin Paldam: The strategies of economic research - An empirical study

2018-05: Ingo Geishecker, Philipp J.H. Schröder, and Allan Sørensen: One-off Export Events

2018-06: Jesper Bagger, Mads Hejlesen, Kazuhiko Sumiya and Rune Vejlin: Income Taxation and the Equilibrium Allocation of Labor

$\hat{T}(\cdot)$, but holding the marginal utility of consumption constant at the pre-reform level, i.e. at $u(c(a))$. This counterfactual object is defined by the functional

$$\zeta'(\tilde{s}(w|a)) = \lambda(a) \int_{w}^{\overline{w}(a)} \frac{u'(c(a))[1 - \hat{T}'(x)]\overline{F}(x|a)}{\rho + \delta(a) + \lambda(a)\mu + \lambda(a)\tilde{s}(x|a)\overline{F}(x|a)} dx. \tag{18}$$

The full partial equilibrium on-the-job search effort response of a type-a worker employed with wage w, $\hat{s}(w|a) - s(w|a)$ can be decomposed into a substitution effect and an income effect according to

$$\hat{s}(w|a) - s(w|a) = \underbrace{\tilde{s}(w|a) - s(w|a)}_{\text{Substitution effect}} + \underbrace{\hat{s}(w|a) - \tilde{s}(w|a)}_{\text{Income effect}}. \tag{19}$$

To fix ideas, suppose the tax reform increases the net-of-tax rate for wages in the interval $[\hat{w}, \hat{w} + \Delta] \in [b, \overline{w}(a)]$ for some $\Delta > 0$, while leaving the rest of the tax function unchanged. Suppose $w < \hat{w}$. Then, inspection of (13) and (18) reveals that, in response to the higher net-of-tax rates in $[\hat{w}, \hat{w} + \Delta]$, on-the-job search effort at wage $w < \hat{w}$ increases, i.e the substitution effect $\tilde{s}(w|a) - s(w|a)$ is positive. The magnitude of the substitution effect is inversely related to the convexity of the disutility from search effort, i.e. $\zeta''(\cdot)$. The intuition is straightforward: An increase in the net-of-tax rate at some interval $[\hat{w}, \hat{w} + \Delta] \in (w, \overline{w}(a)]$ increases the gains from reallocating to this wage interval, mandating an increase in search effort at wage w to rebalance marginal costs and benefits. If the disutility from search is highly convex, only small adjustments in search effort are needed to equalize marginal costs and benefits. If $w > \hat{w} + \Delta$, the reform affect only wages below the worker's current position, leaving job search incentives unchanged, and the substitution effect is zero.[11]

The increased net-of-tax wages in the interval $[\hat{w}, \hat{w} + \Delta]$ also increases workers' permanent income, and thus consumption. That is, $c(a) < \hat{c}(a)$, and by the convexity of the utility of consumption, $u'(c(a)) \geq u'(\hat{c}(a))$. Inspection of (17) and (18) reveals that the magnitude of the income effect is driven by the curvature of the utility of consumption, $u(\cdot)$. Again, the intuition is straightforward: An increase in the net-of-tax rate at some interval $[\hat{w}, \hat{w} + \Delta]$ increases the worker's permanent net-of-tax income, reducing the marginal utility of consumption and thus mandating a decrease in search effort at all wage levels. Since the income effect operates through the workers' permanent income, it is present even when $[\hat{w}, \hat{w} + \Delta] \in [b, w]$.

2.3 Firm behavior

Firm behavior is modeled in a one-shot game. Conditional on entry, a firm sets a wage policy and a hiring intensity policy in order to maximize its expected flow of net-of-tax profit in steady state. In doing so, each firm takes the income and profit tax regimes, workers' search strategies, and the hiring and wage policies of other firms as given. Each firm conditions its policies on worker ability.[12] Specifically, a wage policy for a generic type-$p \in [p_0, 1]$ firm is

[11] In the case where $w \in [\hat{w}, \hat{w} + \Delta]$ a positive substitution effect arises from the incentives to reallocate to wages in the interval $(w, \hat{w} + \Delta]$.

[12] This setup involves a degree of directedness in firms' search processes, eliminating cross-ability congestion externalities in wage policies.

$w(a|p) : [0,1] \mapsto \mathbb{R}^+$, and a hiring intensity policy is $v(a|p) : [0,1] \mapsto \mathbb{R}^+$. Recall that the productivity distribution among the entering firms is $\Gamma_0(\cdot)$. There is an underlying exogenous distribution of potential productivities $\Gamma(p)$ with $p \in [0,1]$. There is free entry of firms with entry decisions based on expected steady state net-of-tax profit flow. The entry game pins down the measure of entering firms, M_0, the minimum productivity p_0, as well as $\Gamma_0(\cdot)$.

2.3.1 Profit maximization

Following Burdett and Mortensen (1998) and Bontemps, Robin, and van den Berg (1999, 2000), firms that have entered the labor market seek to maximize the expected steady state flow profit.[13] Let $\ell(a|w,v)$ be the the steady state measure of type-a workers that a generic (type-p) firm can expect to employ when it offers this type of workers a wage w while searching for them at intensity v. The notation makes it clear that firm productivity p plays no direct role in shaping $\ell(a|w,v)$, but, of course, the optimal wage and hiring policies will depend on p, because output per worker depends on p. Indeed, profit per type-a worker in a type-p firm, paying a wage w is $y(a,p) - w$. Let $\pi(p)$ be the steady state net-of-tax profit flow resulting from a type-p firms' optimal wage and hiring intensity policy. Recall that profit flows are taxed at a flat rate τ. It follows that

$$\pi(p) = \max_{w(\cdot|p),v(\cdot|p)} (1-\tau) \int_0^1 \Big\langle [y(a,p) - w(a|p)]\ell(a|w(a|p), v(a|p)) - d(v(a|p)) \Big\rangle da, \tag{20}$$

where $d : [0,\infty) \mapsto [0,\infty)$ is the twice continuously differentiable hiring intensity cost function. We impose $d'(\cdot) > 0$ and $d''(\cdot) > 0$. Equation (20) embeds a number of assumptions stated above. First, firm-level output exhibits constant returns to scale in matches. Second, hiring cost is a monetary cost that is fully deductible in relation to profit taxation (say, the cost of running a human resource department). In the current formulation of the model, the hiring costs vanishes from the economy, i.e. it does not accrue to any of the agents in the economy.[14] We are currently working on a version of the model where firms assign employees to either production or recruitment activities, similar to the formulation in Shimer (2010). Production workers produce output according to a production function as above, whereas recruitment workers enable the firm to attract more workers. In such a setup, hiring cost can be regarded literally as wage costs in a human resource department, and are clearly accruing to the (recruitment) workers. Finally, a third assumption embedded in (20) is that hiring intensity can be directed towards workers of a particular ability, with a firms' total hiring cost being the integrated worker-type specific hiring costs.

In order for a firm to be able to hire any type-a workers at all, its wage offer must exceed the type-a worker reservation wage and the firm must exert strictly positive hiring effort. That is, $\ell(a|w,v) = 0$ if and only if $w < b$ or $v = 0$. Consider a generic type-p firm with type-a worker

[13] Burdett and Mortensen (1998) and Mortensen (2005) point out that steady-state profit is the appropriate criterion only if firms do not discount the future, namely $\rho = 0$.

[14] A contrived and empirically nonsatisfactory narrative underlying such a formulation could involve firms outsourcing their hiring operation to foreign owned recruitment agencies.

wage and hiring policies represented by $w(a|p) > b$ and $v(a|p) > 0$. Then, $\ell(a|w(a|p), v(a|p))$ is pinned down by the following balanced flow equation,

$$\ell(a|w(a|p), v(a|p))\left[\delta(a) + \lambda(a)\mu + \lambda(a)s(w(a|p)|a)\overline{F}(w(a|p)|a)\right] = \frac{v(a|p)}{V(a)}\Big\langle h(a)n^0(a)[\lambda(a)\mu + \lambda(a)s(b|a)] + h(a)[1 - n^0(a)]\left[\lambda\mu + \lambda(a)\int_b^{w(a|p)} s(x|a)dG(x|a)\right]\Big\rangle, \quad (21)$$

where $s(w|a)$ is given by (13), $n^0(a)$ is the steady state type-a worker unemployment rate (14), $G(w|a)$ is given by (16), and where $V(a) = M_0 \int_{p_0}^1 v(a|p)d\Gamma_0(p)$ is the integrated type-a worker hiring intensity. The left-hand side of (21) is the measure of type-a workers that leave a type-p firm with type-a worker wage and hiring policies $w(a|p)$ and $v(a|p)$. These workers leave either because they are laid off, reallocated or because they find a higher paying job. The right-hand side of (21) is the measure of type-a workers that are hired into the type-p firm. The expression in angle brackets is the number of meetings between firms and type-a workers willing to accept a job offer $w(a|p)$. These workers are either coming from unemployment, reallocated from other firms, or are currently working in firms paying a wage less than $w(a|p)$. The pool of willing workers are distributed across hiring firms in proportion to the firms' hiring intensity, i.e. according to the density $v(a|p)/V(a)$.

Since (21) implies that $\ell(a|w(a|p), v(a|p)) = \frac{v(a|p)}{V(a)}\tilde{\ell}(a|w(a|p))h(a)$, where

$$\tilde{\ell}(a|w(a|p)) = \frac{n^0(a)[\lambda(a)\mu + \lambda(a)s(b|a)] + [1 - n^0(a)][\lambda(a)\mu + \lambda(a)\int_b^{w(a|p)} s(x|a)dG(x|a)]}{\delta(a) + \lambda(a)\mu + \lambda(a)s(w(a|p)|a)\overline{F}(w(a|p)|a)}, \quad (22)$$

we can analyze a firm's wage policy $w(a|p)$ independently of its hiring intensity policy $v(a|p)$. Lemma 2 states a useful and intuitive property shared by $\ell(a|w(a|p))$ and $\tilde{\ell}(a|w(a|p))$.

Lemma 2. *$\ell(a|w)$ and $\tilde{\ell}(a|w)$ are strictly increasing in w for $w \in [b, \overline{w}(a)]$.*

Proof. See Appendix A.5. □

The intuition behind Lemma 2 is straightforward. By offering a higher wage, a firm simultaneously increases its ability to poach workers from other (lower paying) firms, and reduces the likelihood of its own workers being poached by other (higher paying) firms. Both effects tend to increase the steady state measure of type-a workers employed in the firm.

Using (21) and (22) allow us to restate the firms profit maximization problem (20) as a two-step problem. Indeed,

$$\pi(p) = \max_{v(\cdot|p)}(1 - \tau)\int_0^1 \Big\langle \frac{v(a|p)}{V(a)}\tilde{\pi}(a|p)h(a) - d(v(a|p))\Big\rangle da, \quad (23)$$

where

$$\tilde{\pi}(a|p) = \max_w \Big\langle [y(a, p) - w]\tilde{\ell}(a|w)\Big\rangle, \quad (24)$$

with $\tilde{\ell}(a|w)$ given by (22). Equation (24) pins down the optimal wage policy $w(\cdot|p)$, independently of the hiring policy $v(\cdot|p)$. Equation (23) then pins down the optimal hiring policy $v(\cdot|p)$, given the optimal wage policy.

The fact that hiring costs are fully deductable ensures that profit taxation has no direct distortionary effects on hiring intensity, nor on wage policies. In equilibrium, profit taxation will, however, have distortionary effects through its impact on firm entry and exit, and the derived effects of this on firms' wage policies. As established above, see (13), income taxation has distortionary effects on workers' search effort and thus on the allocation of labor, see (14) and (16). It is evident from (22), (23) and (24) that income taxation therefore induce further equilibrium distortions of firms' wage and hiring policies.

2.3.2 Optimal wage and hiring policies

We restrict attention to a pure strategy equilibrium in which the solution to (23) and (24) are two functions $w : [0,1] \times [p_0, 1] \mapsto \mathbb{R}$ and $v : [0,1] \times [p_0, 1] \mapsto \mathbb{R}$ mapping worker-types and firm-types into wage offers and hiring intensities. In a similar, but simpler, model where search effort and hiring intensity are exogenously fixed at constant levels, Bontemps, Robin, and van den Berg (2000) show that, provided that the underlying productivity distribution $\Gamma_0(\cdot)$ is continuous, the pure strategy equilibrium is in fact the only equilibrium. We conjecture that this result carries over to our slightly more complicated setup.

The optimal wage and hiring intensity policies, $w(a|p)$ and $v(a|p)$, have a couple of useful and intuitive properties that are formally stated below.

Proposition 1. *$w(a|p)$ is strictly increasing in p for all $a \in [0,1]$.*

Proof. See Appendix A.6. □

Proposition 2. *$v(a|p)$ is unique and is strictly increasing in p for all $a \in [0,1]$.*

Proof. See Appendix A.7. □

Proposition 3. *$\pi(p)$ is strictly increasing in p.*

Proof. See Appendix A.8. □

2.3.3 Firm entry and exit

Firm entry and exit is as in Melitz (2003). Firms enter the market by paying a sunk entry cost C. Our current model formulation assumes the presence of a third class of economic agent in the economy besides workers and firms, namely entrepreneurs. An entrepreneur is a risk neutral agent who therefore do not participate in the risk sharing arrangement in the large households. An entering firm is an entrepreneur who pays the sunk cost to develop a business idea. One might for example think of C as a money-metric disutility cost of an entrepreneurs foregone leisure. Indeed, we treat C as non-deductible in relation to profit taxation. Immediately upon entry, the firm observes its productivity p, and decides whether to hire workers and start producing, or exit. A firm's productivity p is drawn from the exogenous productivity distribution $\Gamma(\cdot)$ with support $[0,1]$. A firm with productivity p exits if the steady profit at that productivity is negative, i.e. if $\pi(p) < 0$. Proposition 3 established that $\pi(p)$ is strictly increasing in p. Hence, there is a threshold productivity p_0 such that firms with $p < p_0$ exit the market. The threshold productivity p_0 solves $\pi(p_0) = 0$ and the endogenous distribution of productivities across operating firms, $\Gamma_0(\cdot)$, is given by

$$\Gamma_0(p) = \frac{\Gamma(p) - \Gamma(p_0)}{\overline{\Gamma}(p_0)}, \quad p \in [p_0, 1]. \tag{25}$$

Free entry ensures that, in equilibrium, the measure of active firms in the economy, M_0, adjust such that the expected steady state profit flow from entry is offset by the opportunity (flow) cost of entry, given by the annuitized entry cost ρC. That is,

$$\int_{p_0}^{1} \pi(p) d\Gamma(p) = \rho C. \tag{26}$$

Substituting (23) for $\pi(p)$ in (26), and using that $V(a) = M_0 \int_{p_0}^{\overline{p}} v(a|x) d\Gamma_0(x)$, implies that

$$M_0 = \left[\frac{\rho C}{1-\tau} + \int_{p_0}^{1}\int_{0}^{1} d(v(a|p)) da d\Gamma(p)\right]^{-1} \left[\int_{p_0}^{1}\int_{0}^{1} \frac{v(a|p)\tilde{\pi}(a|p)h(a)}{\int_{p_0}^{1} v(a|x) d\Gamma_0(x)} da d\Gamma(p)\right]. \tag{27}$$

Equation (27) shows that profit taxation distorts firms' entry and exit decisions, and thus impact the measure of active firms. These distortion arise because entry costs are sunk, and therefore not deductible. Profit taxation effectively increases the cost of entry, or reduces the value of entry, thereby depressing the number of firms in the economy.

Ownership structure of firms. Firms are owned by the risk neutral entrepreneurs who simply consumes the net-of-tax profit flow, on average ρC. As mentioned above, the entrepreneurs do not participate in the risk sharing arrangement of the risk averse workers. Indeed, the household budget constraint (4) does not include any net-of-tax profits. We stress that the entrepreneurs are within the jurisdiction of the tax authorities as the revenue from profit tax enters the government's budget (1).

An alternative model formulation assumes that the households are the ultimate owners of the firms, perhaps through a mutual fund. In such a formulation, net-of-tax profit flows accrue to households and enters the household budget constraint (4), and will have implications for workers' search behavior via income effects. While such a model formulation is perhaps more conventional in the macro literature, we note that the vast majority of individuals in our data in fact have negative capital income, see Figure 1.[15] Hence, from a purely empirical standpoint, it is not clear that capital income generates significant income effects in workers search behavior. Nonetheless, to explore these issues further we are currently working on a model formulation where household own the firms through mutual funds.

[15] This result arise because the measured capital income include interest on mortgage debt. Taking out mortgage interest, we believe most individuals in our data have little or not capital income.

2.3.4 Aggregation

Since firms optimal worker-type a conditional wage offers $w(a|p)$ are strictly increasing in firm-type p, see Proposition 1, and firms meet potential employees at a rate proportional to hiring intensity $v(a|p)$, the wage offer distribution faced by workers is simply the hiring intensity weighted productivity distribution given by

$$F(w|a) \equiv F(w(a|p)|a) = \frac{\int_{p_0}^{p} v(a|x)d\Gamma_0(x)}{\int_{p_0}^{1} v(a|x)d\Gamma_0(x)}. \tag{28}$$

Let $\theta(a)$ be the worker-type a specific labor market tightness, measured as the ratio of aggregate hiring intensity $V(a) = M_0 \int_{p_0}^{1} v(p)d\Gamma_0(p)$ to aggregate search effort $S(a)$ in the worker-type a labor market. Aggregate worker-type a search effort $S(a)$ is given by

$$S(a) = n^0(a)[\mu + s(b|a)] + [1 - n^0(a)] \int_b^{\overline{w}(a)} [\mu + s(w|a)]dG(w|a), \tag{29}$$

where $s(w|a)$ is given by (13), $n^0(a)$ is given by (14), and $G(w|a)$ is given by (16), and where, by Proposition 1, $\overline{w}(a) = w(a|1)$, the optimal wage offer in the most productive firm with $p = 1$.

The model structure implies that the flow of contacts between type-a workers and firms in the economy is $\lambda S(a)$. With a constant returns to scale Cobb-Douglas matching function with elasticity of matches with respect to unemployed search being η, we have that $\lambda(a) = \theta(a)^{1-\eta}$. This closes the model, and we can state the equilibrium conditions in the next subsection.

2.4 Equilibrium

Our equilibrium definition follows Mortensen (2005).

Definition 1. *A steady-state labor market equilibrium is composed of the following components:*

- *The optimal reservation strategy: An employed worker accepts any job offer that offers a wage strictly greater than the current wage. An unemployed type-a worker accepts any job offer greater than or equal to $\phi(a) = b$, see (12).*
- *Workers' optimal search effort solves (13).*
- *Firms' worker-type specific optimal wage and hiring intensity policies $w(a|p)$ and $v(a|p)$ solve (24) and (23).*
- *The worker-type specific unemployment rate $n^0(a)$ satisfies (14) and the distribution of wages across employed workers $G(w|a)$ satisfies (16).*
- *The conditional gross wage offer distribution $F(w|a)$ that satisfies (28).*
- *The threshold productivity for entry p_0 solves $\pi(p_0) = 0$ with $\pi(p)$ given by (20).*
- *The measure of firms operating in the market M_0 given by (27).*

- *The worker-type specific matching efficiency parameters* $\lambda(a)$ *are given by* $\lambda(a) = \theta(a)^{1-\eta}$ *where* $\theta(a) = S(a)/V(a)$, $V(a) = M_0 \int_{p_0}^{1} v(p) d\Gamma_0(p)$ *and* $S(a)$ *given by* (29).

Given the rich structure of the model, we have not been able to prove existence of the equilibrium. However, our solution algorithm finds a numerical solution under reasonable parameter values. The numerical model solution is detailed in Appendix D.

3 Data

Our empirical analysis uses a matched employer-employee panel covering the entire Danish population during 1990-2005. The data is constructed from administrative records. Estimation of the structural parameters of the model is carried out using a subset of this data pertaining to the period 1999-2003. Evaluation of a series of tax reforms implemented in 1994, 1998, and 2004 requires the use of the entire 1990-2005 data period.

3.1 Data sources

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

Labor market spell data. The labor market spell data contains individual job and nonemployment 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.[16] We restrict attention to the period 1990-2005. A job spell is defined as a continuous period of primary employment at a given firm, with duration measured in days.[17] Job spells, and therefore labor market transitions, are recorded at the firm-level. A firm may consist of multiple workplaces, but the labor market spell data is constructed such that continuous employment at different workplaces within a firm is considered as a single job spell. Nonemployment spells are periods where no job spells are recorded. We are not able to distinguish between nonemployed who actively search for a job and nonparticipants.

The unit of observation in the labor market spell data is a person-spell-year. Hence, a spell that starts on Jan. 5th, 2000 and ends on March 12th, 2002 is represented by three observations relating to the years 2000, 2001 and 2002. The job spell data includes worker and firm identifiers, start- and end-dates of the job, the annual earnings pertaining to the job, as

[16] Henning Bunzel at Aarhus University has been instrumental in developing the job spell data described here.

[17] 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.

well as an estimate of the annual number of hours worked in the job.[18] Hence, we can compute (an estimate of) the average annual wage rate in each job. In cases where a worker transitions from one firm to another with up to 14 days of nonemployment between the two employment spells, we overwrite the the nonemployment spell and thus record a job-to-job transition. In cases where a worker ends an employment spell at a firm, but transitions back to the same firm with at most 12 weeks of non-employment between the two employment spells, we overwrite the nonemployment spell and instead record a single employment spell.

IDA data. IDA consists of several sub-panels. We use the sub-panels IDA-P and IDA-S from the period 1990-2005. IDA-P contains annual background information from public registers on all individuals aged 15-74 residing legally in Denmark on the 31st of December. IDA-P information includes highest completed education including date of completion, and information on any ongoing education subsequent to the highest completed education. We use date of completion of the highest completed education is used to define labor market entry. The unit of observation in IDA-P is a person-year. IDA-S contains background information from public records on all physical workplaces in Denmark. We retain a public sector indicator from IDA-S, defining a firm's public sector status to be the public sector status of its largest workplace. The unit of observation in the (aggregated) IDA-S panel is a firm-year.

VAT data. Data on firms' (not workplaces) sales and purchases are obtains from the panels MOMS and MOMM, constructed from firm's value added tax (VAT) accounts maintained by the Danish tax authorities. Danish firms settle VAT accounts monthly, quarterly or annually depending on the size of their revenue. MOMS covers the period 1995-2000 and contains annual sales and purchases. MOMM is a monthly panel starting in January, 2001. We use MOMM data up until December, 2005, aggregating the monthly information to an annual frequency.[19] Putting MOMS and MOMM together, we obtain an annual panel of firms' sales and purchases for the period 1995-2005. 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. Merging the data sources is relatively straightforward using the available person and firm identifiers, and Appendix B provide a detailed description of the merging procedure. Here we simply note that the resulting 1990-2005 matched employer-employee panel contains 84,474,045 spell observations on 4,447,401 individuals, 428,448 firms, 57,568,393 job spells and 26,905,652 non-employment spells prior to selection of the analysis

[18] Annual hours are estimated using information on mandatory pension contributions. Lund and Vejlin (2015) develop and implement a procedure for computing annual hours in a job in the IDA data for the period 1980-2007, primarily using information on mandatory pension contributions. We adapt this procedure for the spell data with some minor simplifications.

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

data.

3.2 Selecting the analysis data

We start by discarding all observations on individuals never observed with either age or education information.[20] We also discard individuals with implausible education information, defined as age minus years of schooling being less than 5. We then define labor market entry to occur at the observed date of graduation from highest completed education (as observed in the window of observation, 1990-2005), or at January, 1st of the year an individual turns 19, whichever occur at the latest date. All pre-entry observations are discarded. We define labor market exit to occur at the last year an individual is observed residing legally in Denmark (i.e. present in IDA-P) or at December, 31st of the year an individual turns 55, whichever occur first. All post-exit observations are discarded. We right censor observations at age 55 to avoid labor market behavior driven by retirement considerations.

A series tax reforms were implemented in Denmark on January 1st in 1990, 1994, 1999 and 2004. As we document further below, between these reform dates, the tax system is relatively stable, and we therefore extract four short matched employer-employee panels pertaining to the periods 1990-1993, 1994-1998, 1999-2003, and 2004-2005. We employ a rudimentary, but consistent, set of selection criteria to each of the four panels. These criteria are imposed independently for each panel. Hence, an individual or firm discarded from, say, the 1999-2003 panel due to violation of the selection criteria to be outlined below, will be included in the other three analysis datasets if the same selection criteria are not violated during in the 1990-1993, 1994-1998 or 2004-2005 panels.

Effectively, we select individuals who primarily work full time in the private sector (when they work). Specifically, for each of the four panels covering 1990-1993, 1994-1998, 1999-2003, and 2004-2005, we discard workers who, (i) in any year worked less than 25 hours a week on average, when working, (ii) at any point in time work in the public sector, or (iii) leave the panel prematurely (likely caused by periods of living abroad, or death). Finally, for the estimation period 1999-2003, we trend value added and wages to 2003 prices using the internal wage deflator.

The selection process leaves us with four short matched employer-employee panels for the periods 1990-1993, 1994-1998, 1999-2003, and 2004-2005. Table 1 provides basic summary statistics for the estimation period 1999-2003. We note that the number of workers and firms, as well as the non-employment rates is relatively stable. Table 2 report some further basic descriptive statistics on the estimation dataset. On average, a worker is observed for slightly more than 4 years, during which she works at 1.84 firms. The descriptive statistics on age and education is in line with aggregate statistics for the Danish labor force.

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

Table 1: November 28th cross-sections in the 1999-2003 estimation panel

Year	# Persons	# Firms	Nonemployment rate
1999	932,868	75,310	28.9%
2000	895,223	71,710	28.8%
2001	861,383	71,890	28.6%
2002	831,835	70,423	28.7%
2003	813,555	70,902	30.3%

Table 2: Summery Statistics for the 1999-2003 estimation panel

	Mean	Std. Dev.	Percent
Number of years in sample	4.34	1.29	
Number of workers per firm	8.53	93.16	
Number of firms per worker	1.84	1.24	
Age	41.30	8.93	
Highest completed education			
Primary			36.75
High school			6.06
Vocational education			44.38
Short tertiary education			3.20
Medium tertiary education			5.86
Bachelor			0.57
Master degree or longer			3.17

Note: Numbers of years in the sample is computed as the number of years (maximum five) a given worker is observed in the sample. Number of workers per firm is the number of workers a given firm employs at a given November cross section. Number of firms per worker is the number of firms a given worker is observed with during the panel.

4 The empirical tax function

The income tax function is a key component in our analysis. This section provides a brief overview of the Danish income tax system, discusses the 1994, 1999, and 2004 tax reforms, and describe how we estimate marginal and average tax functions using the micro-level analysis panels described above.

4.1 The Danish income tax system 1990-2005

An overview. Income taxes in Denmark are divided into regional and national taxes. Regional taxes include municipality tax, county tax, and a minuscule church tax paid only by members of the Church of Denmark. Regional tax rates differ by municipality and county, although the variation in tax rates is small. National taxes have a progressive structure with three brackets (bottom, middle and top) with tax rates cumulated across brackets. Regional and national taxes differ not only in tax rates, but also in terms of the base on which they are levied. Furthermore, during the period we consider, various social security and labor market contributions (SSC and LMC), as well as earned income tax credits (EITCs) have been introduced (and in some cases abolished again).[21] Finally, a marginal tax ceiling is in place. If the sum of all tax rates, excluding the church tax, and any social and labor market contributions, exceeds this ceiling, the top tax rate is adjusted to equal the ceiling. The Danish tax code is based on individual filing for married couples, but some components of deductions or allowances can be transferred across spouses. We take the spousal income into account when we simulate marginal tax rates.

The tax base of regional and national income taxes consists of a combination of income concepts, including personal income, capital income, and deductions. Table 3 details the main items in each of the taxable income components relevant for the taxation of labor income during 1990-2005, and Table 4 provides descriptive statistics on the distribution of each of these components in the 1999-2003 estimation data.[22] We note that, on average, labor income dwarfs the other income types, as well as the deductions, but also that there are considerable variation in each of the taxable income components, highlighting the benefit of having comprehensive micro data at our disposal. Finally, capital income is negative for the majority of Danish taxpayers as a result of interest payments on mortgages and loans.

To get a sense of the correlation structure between the different income concepts, Figure 1 plots the average personal income net of labor income, capital income and deductions against

[21] Social security contributions are compulsory payments that confer entitlement to future social benefit, including unemployment benefits. For example, a labor market contribution introduced in 1994, was at the onset effectively a social security contribution. Earmarking of the labor market contributions ceased in 2008, after which it is best described as a regular (flat) income tax.

[22] Since 1987, the taxation of wealth including income from stock holdings is completely separate from the taxation of labor income. Consistent with the theoretical model our empirical analysis ignores the role of wealth in shaping labor market behavior. See Jakobsen, Jakobsen, Kleven, and Zucman (2018) for a comprehensive analysis of wealth taxation in Denmark.

Table 3: Income concepts in the Danish income tax system

Income concept	Main items included
Labor income (LI)	Salary, wages, honoraria, fees, bonuses, fringe benefits, business earnings
Personal income (PI)	$LI+$ transfers, grants, awards, gifts, received alimony $-$ labor market contribution, certain pension contributions
Capital income (CI)	Interest income, rental income, business capital income $-$ interest on debt (mortgages, bank loans, credit card debt, student loans)
Deductions (D)	Commuting, union fees, UI contributions, other work expenditures, charity, paid alimony

Note: Reproduction of Table 1 in Kleven and Schultz (2014).

Table 4: The summary statistics on the distribution of 1999-2003 taxable income components

	Average	Std. dev.	P25	P50	P75
Labor income	333,763	184,862	245,640	301,755	381,665
Personal income net of labor income	4,485	590,762	$-3,752$	1	2,001
Capital income	$-25,158$	99,442	$-41,549$	$-22,281$	$-4,288$
Deductions	16,691	13,579	9,489	13,461	19,719

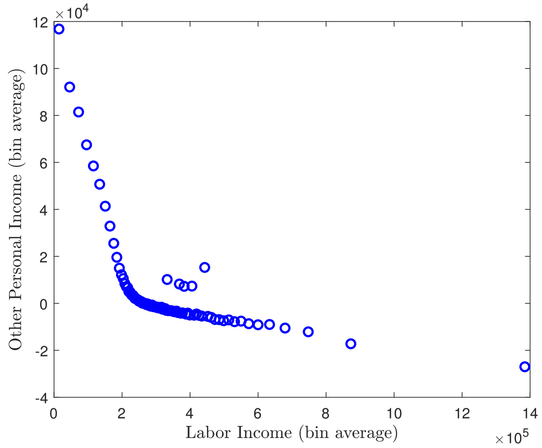
Note: All amounts denominated in 2003 Danish Kroner. P25, P50 and P75 refers to the 25th, 50th and 75th percentiles in the distribution of each of the variables (pooled across the years 1999-2003) the percentiles in the distribution of labor income for the 1999-2003 panel used in our estimation. From panel A, we see that, naturally, personal income net of labor income is, on average, higher for individual with low labor income, reflecting primarily transfers and benefits related to periods of nonemployment. However, once labor income reaches about 250,000 Danish Kroner—around the 25th percentile in the labor income distribution—the personal income net of labor income component is, on average, close to zero, and remains smal as we move through the labor income distribution. In fact, personal income net of labor income becomes, on average, negative for high labor incomes, most likely reflecting certain pension contributions. Panel B in Figure 1 confirms that capital income is, on average, negative for all labor incomes, and that capital income is negatively related to the labor income component. This relationship reflects that, for most individuals, capital income consists almost exclusively of interest on mortgage debts, and that individuals with higher labor income tend to have larger mortgages. Still, the ratio of capital income to labor income remains small throughout the distribution of labor incomes. Finally, Panel C in 1 shows that deductions are a hump-shaped function of labor income, and that the amounts deducted are small relative to labor income throughout the distribution of labor incomes.

Overall, the descriptive evidence in Table 4 and in Figure 1 clearly indicate that labor income is the primary source of income for almost all individuals in our data, and that, for most individuals, capital income and deductions will have not greatly affect their available permanent income. We will account for the presence of non-labor income components in the estimation of the tax function detailed below, but our model ignores any consumption enjoyed out of non-labor income. The evidence in Table 4 and in Figure 1 suggest that the bias induced by this omission is likely to be small.

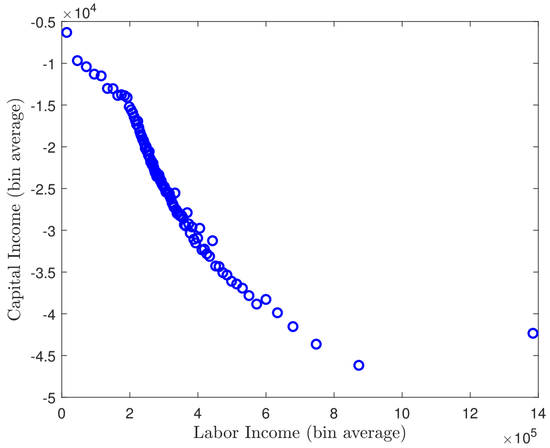
Income tax reforms. There were three major reforms of the income tax system in Denmark during 1990-2005, namely in 1994, 1999, and 2004.[23] Although some of the reforms were part of a larger fiscal reform packages, all three reforms were intended to increase the return to work. The reforms implemented changes to the tax bases of regional and national taxes, as well as the tax rates, and as we shall see, induced significant shifts in the marginal income tax function faced by individuals. Table 5 describes the tax bases and rates in place during the periods 1990-1993, 1994-1998, 1999-2003, and 2004-2005. We note that the three reforms were characterized by (i) changes in tax bases in all national tax brackets, (ii) differential changes in progressive tax rates across brackets, (iii) abolition of social security contributions and introduction of labor market contributions and EITCs, and (iv) rising regional tax rates and declining marginal tax ceilings. We show further below that the the income levels defining bottom, middle and top tax brackets also change during the period (over and above indexation to inflation).

[23] The description of the reforms to come is based on material available on the website of the Danish Ministry of Taxation, see http://www.skm.dk/aktuelt/temaer/skattereformer-og-skatteomlaegninger-siden-1987 (in Danish).

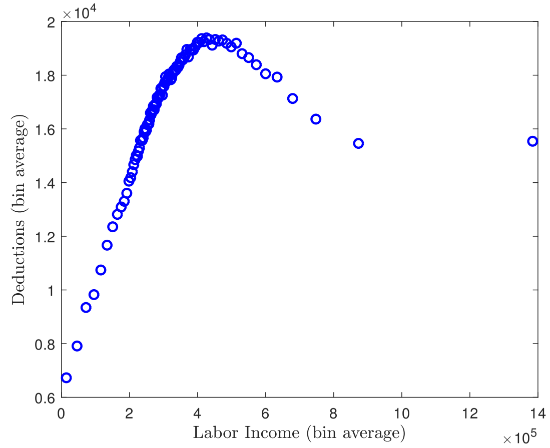
Figure 1: Average personal income, capital income, and deductions by labor income percentiles



Panel A: Personal income net of labor income



Panel B: Capital income



Panel C: Deductions

Note: Each of the figures plot averages computed within the percentiles of the distribution of labor income in the pooled 1999-2003 cross sections agains the percentiles of the labor income distribution. All amounts denominated in 2003 Danish Kroner.

Table 5: Danish income tax bases and rates 1990-2005

	1993		1998		2001		2005	
Tax type	Base	Rate (%)	Base	Rate (%)	Base	Rate (%)	Base	Rate (%)
Regional taxes	PI + CI – D	30.2 (avg)	PI + CI – D	32.4 (avg)	PI + CI – D	33.2 (avg)	PI + CI – D	33.3 (avg)
National taxes								
Bottom bracket	PI + CI – D	22.0	PI + CI – D	8.0	PI + [CI > 0] – D	6.25	PI + [CI > 0]	5.5
Middle bracket	PI + [CI > 0]	6.0	PI + CI	6.0	PI + [CI > 0]	6.0	PI + [CI > 0]	6.0
Top bracket	PI	12.0	PI + [CI > 21k]	15.0	PI + [CI > 0]	15.0	PI + [CI > 0]	15.0
SSC	–	–	–	–	–	–	–	–
LMC	–	–	LI	8.0	LI	8.0	LI	8.0
EITC	–	–	–	–	–	–	LI	2.5
Marginal tax ceiling		68.0		58.0		59.0		59.0

Note: Tax rates are cumulative. For example, an individual in the middle bracket in 1992 faces a marginal tax rate on labor income of 33.20% + 6.25% + 6.00% = 45.45%. The 1994 tax reform was gradually implemented over the period 1994-1998. The tax bases and rates reported for 1994-1998 are those in place in 1998 when the reform was fully implemented. SSC is an acronym for Social security contributions. LMC is an acronym for Labor market contributions. EITC is an acronym for Earned income tax credits. For the tax base acronyms, see Table 3.

The 1994 tax reform. The 1994 reform was legislated in 1993 with implementation starting January 1st, 1994, but not completed until 1998. Following recommendation from a commission of appointed experts published in 1992, the purpose of the 1994 tax reform was to improve incentives to work, and to reduce opportunities for tax avoidance. While the 1994 reform largely followed the 1992 expert recommendations, it was considered key element in the the newly formed centre-left government attempts to stimulate the Danish economy. Besides the changes to the income tax bases and rates described in Table 5, the reform made social pensions and cash benefits tax liable to facilitate comparison to earned income. Finally, the 1994 reform involved a reduction in property taxes, and an increase in the use of levies and fees in environmental regulation.

The 1999 tax reform. The 1999 income tax reform was an integral part of wider set of economic reforms known as the "Pentecost reforms" (in Danish: "Pinsepakken") due to the date it was proposed. The reform was legislated in June 1998, with the income tax reforms implemented January 1st, 1999. The reformation of the tax system was motivated by a political desire within the centre-left government to lower tax rates on low incomes, and to increase the marginal tax ceiling. Besides the changes to the income tax bases and rates described in Table 5, the "Pentecost reforms" made changes to property taxation, and further promoted the use of levies and fees in environmental regulation as a source of public revenue.

The 2004 tax reform. The 2004 income tax reform was designed by a centre-right government that had come to power in November 2001. The reform was motivated by desire to incentivize labor market participation and work by lowering the tax on labor income. Compared to the 1994 and 1999 reforms, the changes to the marginal tax rates on labor income was not offset by increases in other fees or levies due to the centre-right government having committed to not increase any taxes, levies or fees upon taking office. Besides the changes to the income tax bases and rates described in Table 5, the 2004 tax reform made small changes to vehicle excise duties, and committed public funds to social care and health care, particularly focused on the elderly.

4.2 Estimating the marginal tax functions

Simulating individual marginal labor income tax rates. Table 5 describes how an individual's marginal tax rate depend on several factors, including the distribution of income across the income concepts listed in Table 3, region of residence, marital status, and, if cohabiting or married, partner income. Fortunately, our data contains all the information necessary to simulate the marginal tax rates faced by individuals in our data. The simulations are carried out using a simulator of the Danish tax system constructed by le Maire and Schjerring (2013) and Kleven and Schultz (2014).[24]

[24] We are grateful to Henrik Kleven and Esben Schultz for making the tax simulator available. The tax simulator consists of a set of SAS-programs which we downloaded from the website of the *American Economic Journal:*

The tax simulator allow us to compute the total tax liability for each individual, for each year that we consider (i.e. 1990-2005). For the purpose of this paper, where we focus on labor income taxation, it is useful to represent individual i's tax liability in year t as $T_t(LI_{it}, \mathbf{z}_{it})$, where $T_t(\cdot)$ is the tax function in year t according to the tax simulator, LI_{it} is individual labor income, and $\mathbf{z}_{it}$ is a vector of other income concepts relevant for the tax liabilities in year t, see Table 3. Following Kleven and Schultz (2014) we compute individual marginal tax rates as

$$T'_t(LI_{it}, \mathbf{z}_{it}) = \frac{T_t(LI_{it} + 100, \mathbf{z}_{it}) - T_t(LI_{it}, \mathbf{z}_{it})}{100}.$$

That is, we add 100 Danish Kroner (approximately 15 US Dollars in 2003) to annual labor income LI, and record the share of this marginal increase that is appropriated by taxes. We stress that marginal tax rates are computed individual by individual, year by year, and reflects individual circumstances such as the composition of income across the income concepts Table 3, region of residence, and relevant household characteristics. Hence, the simulated marginal tax rates represent the actual marginal tax rate faced by each individual in our data.

Estimating the marginal labor income tax function. We use the annual individual marginal tax rates as data points for estimation of a marginal labor income tax function for each tax regime, i.e. for each of the periods 1990-1993, 1994-1998, 1999-2003, and 2004-2005. As mentioned above, the 1994 tax reform was implemented gradually over the period 1994-1998. For that particular tax regime, we therefore estimate a marginal labor income tax function for each year. For the remaining three tax regimes, including our estimation period 1999-2003, we pool the annual marginal tax rates and estimate a single marginal labor income tax function. To facilitate the pooling, we deflate labor income to 2003 levels using the internal deflator from spell data wages described above.

Our structural model of worker search requires a marginal taxes as a function of hourly wages, not annual labor income. We transform the observed labor income into hourly wages by dividing by a measure of annual full-time working hours, as constructed by Lund and Vejlin (2015).[25]

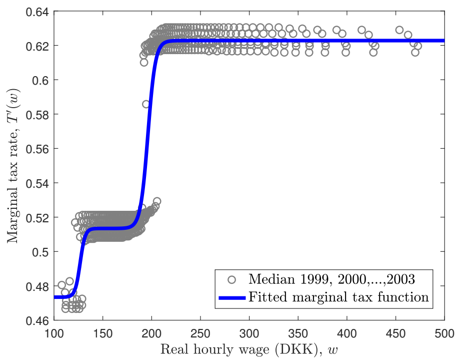
As individual tax liabilities $T_t(LI_{it}, \mathbf{z}_{it})$ are not solely functions of labor income LI, but also depend on other income components, e.g. capital income, and household characteristics, the mapping from hourly wages to marginal tax rates is not one-to-one. To proceed with estimation of the marginal labor income tax functions, we split hourly wages into 100 bins representing the percentiles in the empirical distribution of hourly wages for a given year in a given tax regime. We next compute the median marginal tax rates for each hourly wage bin in each year, pool the annual median marginal tax rates and fit a flexible parametric function (of hourly wages) to these median median marginal tax rates.[26] Figure 2 shows the resulting marginal tax function

Applied Economics. We collected data on regional taxes in Denmark for the period 1985-2005 to use as an input in the tax simulations. This data is available in online supplementary material to this paper.

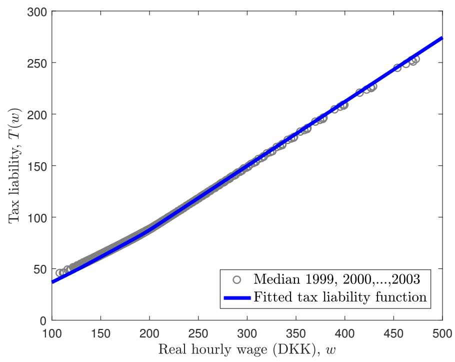
[25] According to this measure, annual hours range between 1,665 and 1,695 for the estimation period 1999-2003.

[26] Appendix C provides further details on this procedure.

Figure 2: Marginal tax function and tax function for Denmark 1999-2003



Panel A: Marginal tax function $T'(\cdot)$



Panel B: Tax function $T(\cdot)$

Note: Panel A plots the marginal tax function $T'(w)$ and Panel B plots the average tax function $T(w)$. The gray circles in Panel A represent the annual within-wage-percentile median tax rates as described in the text. The gray circles in Panel B represent the annual within-wage-percentile median hourly tax liabilities.

(Panel A) and tax function (Panel B) for the estimation period 1999-2003. The marginal tax function in Panel A is superimposed on the annual within-wage-percentile median tax rate, and the average tax function in Panel B is obtained by integrating the marginal tax function. Similar plots of the marginal tax functions and tax functions for the periods 1990-1993, 1994-1998, and 2004-2005 are presented and discussed further below in relation to our tax reform evaluation exercise.

5 Model estimation

Armed with the empirical tax function, we proceed to estimation of the model's structural parameters by way of Indirect Inference, see Gourieroux, Monfort, and Renault (1993). The estimation procedure finds the set of structural parameters which best fit a set of chosen auxiliary statistics. The model is identified if there is one and only one set of structural parameters which generates the auxiliary statistics of the estimation. This section describes how we parameterize the model, discusses issues related to identification of the model parameters, reports and discusses the estimated parameters as well as the model's ability to fit the chosen auxiliary statistics.

5.1 Parameterization

The presentation of the model in section 2 left some key objects, such as the exogenous heterogeneity distributions, match output and workers' felicity function, not fully parametrically specified. For the purpose of estimation, specific parametric forms are required.

Match output $y(a, p)$ is specified as a flexible second order polynomial. Indeed,

$$y(a, p) = \varrho_0 + \varrho_1 a + \varrho_2 p + \varrho_3 a^2 + \varrho_4 p^2 + \varrho_5 ap. \tag{30}$$

where $\boldsymbol{\varrho} = (\varrho_0, ..., \varrho_5)$ are unknown parameters. Consistent with the theoretical model, estimation is carried out under the restriction $y_a(a,p) > 0$ and $y_p(a,p) > 0$.[27] The match output function (30) is sufficiently flexible to warrant a normalization of the exogenous worker and firm heterogeneity distributions to be uniform. That is, $H(\cdot) = \mathscr{U}[0,1]$ and $\Gamma(\cdot) = \mathscr{U}[0,1]$.

Worker felicity (2) has two components, namely utility from consumption, represented by $u(\cdot)$, and disutility from job search effort, represented by $\zeta(\cdot)$. The utility function $u(\cdot)$ is taken to be logarithmic, i.e.

$$u(c) = \log(c), \tag{31}$$

which correspondes to a moderate elasticity of intertemporal substitution. The search effort disutility component is specified as an isoelastic function $\zeta(s)$,

$$\zeta(s - \underline{s}) = \frac{\zeta_0 (s - \underline{s})^{1+1/\zeta_1}}{1 + 1/\zeta_1} \tag{32}$$

which accounts for free search $\underline{s}$, and where $\zeta_0 > 0$ and $\zeta_1 > 0$ are parameters to be estimated. Recall that the curvature of $\zeta(\cdot)$, determined by the parameter ζ_1, is a key determinant of the substitution effect in job search effort.

A brief comment on the restrictive logarithmic specification of the utility function is warranted here. The theoretical analysis made it clear that the curvature of the utility function is a key driver of workers' responses to changes in the tax regimes. We do, however, currently not attempt to estimate this curvature. We are working on adding survey data on consumption expenditures to our data which we conjecture will allow allow us to identify a more general utility function. We are also currently working on establishing whether identification of the curvature in the utility function can be obtained from the already available mobility and wage data. For now, however, the utility function is restricted to be logarithmic, as per (31).

The hiring intensity cost function $\chi(\cdot)$ is specified in an analogous fashion, i.e.

$$\chi(v) = \frac{\chi_0 v^{1+1/\chi_1}}{1 + 1/\chi_1}, \tag{33}$$

where $\chi_0 > 0$ and $\chi_1 > 0$ are parameters to be estimated. The isoelastic search effort disutility and hiring cost functions are standard in the literature, see e.g. Christensen, Lentz, Mortensen, Neumann, and Werwatz (2005), Kreiner, Munch, and Whitta-Jacobsen (2015), and Bagger and Lentz (2015).

For empirical realism we enrich the model to incorporate ability-dependent job destruction rates. Specifically, we let

$$\delta(a) = \exp(\delta_0 + \delta_1 a), \tag{34}$$

and estimate the parameters δ_0 and δ_1. This completes the parameterization of the model.

The fully parameterized model features the following structural parameters $(\rho, b, \boldsymbol{\varrho}, \zeta_0, \zeta_1, \underline{s}, \chi_0, \chi_1, \delta_0, \delta_1, \mu, C, \eta)$. We remind the reader that ρ is the discount rate, b is the unemployment

[27] In addition, we are currently imposing the restriction $y(0,0) = \varrho_0 = b$ in the estimation. Since job search off- and on-the-job are equally efficient, the latter restriction ensures that a match between the least able worker $(a = 0)$ and the least productive firm $(p = 0)$ is viable, and that all workers accept all jobs.

flow income, μ is the reallocation rate, C is the entry cost, and η is the elasticity of the aggregate matching function with respect to job search effort. We fixed a few of the structural parameters prior to estimation. Specifically, we set $\rho = 0.05$ per year, and we also fix the unemployment flow income $b = \varrho_0 = 107$ Danish Kroner (DKK) per hour. In addition, ζ_0 and χ_0 are not separately identified as they both scale the baseline job finding rate λ. We therefore impose the normalization $\chi_0 = 1$, [28] which leaves fifteen structural parameters to be estimated. Collect these in the vector $\boldsymbol{\omega} = (\varrho_1, ..., \varrho_5, \zeta_0, \zeta_1, \underline{s}, \chi_0, \chi_1, \delta_0, \delta_1, \mu, C, \eta)$. Let $\boldsymbol{\omega}_0$ be the true value of $\boldsymbol{\omega}$. The indirect inference estimator is given by

$$\hat{\boldsymbol{\omega}} = \arg\min_{\boldsymbol{\omega}} \left[\mathbf{a}(\boldsymbol{\omega}_0) - \mathbf{a}^S(\boldsymbol{\omega})\right]' \boldsymbol{\Omega} \left[\mathbf{a}(\boldsymbol{\omega}_0) - \mathbf{a}^S(\boldsymbol{\omega})\right], \tag{35}$$

where $\boldsymbol{\Omega}$ is a symmetric and positive definite weighting matrix, and $\mathbf{a}(\boldsymbol{\omega}_0)$ is a vector of auxiliary statistics computed using the data described in section 3, i.e. at the unknown true value of the structural parameter vector $\boldsymbol{\omega}_0$. The vector $\mathbf{a}^S(\boldsymbol{\omega}) = S^{-1}\sum_{s=1}^{S}\mathbf{a}_s(\boldsymbol{\omega})$ contains the same auxiliary statistics, but computed using N^s labor market histories simulated from the structural model for some value $\boldsymbol{\omega}$ of the structural parameter value. Effectively, the Indirect Inference estimator finds the structural parameter vector that minimizes the distance between the real data and data simulated from the structural model, as measured by the included auxiliary statistics, and the weight matrix $\boldsymbol{\Omega}$.[29] Appendix D contains a detailed description of the numerical computation of the model equilibrium, as well as a description of the simulation procedure we apply.

Implementation of (35) requires first of all a set of auxiliary statistics, and we discuss the ones we include below. In addition, however, implementation requires us to set S, the number of simulation repetitions, N^s, the number workers (i.e. the number of simulated labor market histories per simulation repetition s), as well as the weight-matrix $\boldsymbol{\Omega}$ for computing the distance between empirical and simulated auxiliary statistics, see (35). In our implementation, we take $S = 1$, $N^s = 100,000$, and take $\boldsymbol{\Omega}$ to be a diagonal matrix. With a few exceptions reported further below, the diagonal entries in $\boldsymbol{\Omega}$ are set to 1.[30]

[28] The same argument applies to any multiplicative constants in the aggregate matching function, which is also implicitly normalized to unity.

[29] Under mild regularity conditions, see Gourieroux, Monfort, and Renault (1993), we have $\sqrt{N}(\hat{\boldsymbol{\omega}}^S - \boldsymbol{\omega}_0) \xrightarrow{d} \mathcal{N}(\mathbf{0}, (1+S^{-1})\mathbf{V}_0)$, where $\mathbf{V}_0 = [\mathbf{J}_0'\boldsymbol{\Sigma}_0\mathbf{J}_0]^{-1}\mathbf{J}_0'\boldsymbol{\Sigma}_0\boldsymbol{\Omega}_0\boldsymbol{\Sigma}_0\mathbf{J}_0[\mathbf{J}_0'\boldsymbol{\Sigma}_0\mathbf{J}_0]^{-1}$, $\mathbf{J}_0 = \partial\mathbf{a}(\boldsymbol{\omega})/\partial\boldsymbol{\omega}$ at $\boldsymbol{\omega} = \boldsymbol{\omega}_0$, $\boldsymbol{\Sigma}_0$ is the variance-covariance matrix of $\boldsymbol{a}(\boldsymbol{\omega}_0)$. We report standard errors based on the asymptotic variance-covariance matrix $N^{-1}(1+S^{-1})\hat{\mathbf{V}}$, where $\hat{\mathbf{V}}$ is $\mathbf{V}_0$ with true values replaces by estimates. $\boldsymbol{\Sigma}_0$ is estimated by block-bootstrapping $\boldsymbol{a}(\hat{\boldsymbol{\omega}})$, with workers' labor market histories constituting the blocks. We take N to be the number of workers in our estimation data.

[30] Asymptotic efficiency of the Indirect Inference estimator requires $\boldsymbol{\Omega} = \hat{\boldsymbol{\Sigma}}^{-1}$, i.e. using the inverse variance-covariance matrix of the vector of auxiliary statistics as weight-matrix. Altonji and Segal (1996) illustrates how an asymptotically efficient GMM estimator may be marred by small sample biases in applications.

5.2 The government budget

In steady state, a type-a worker's contribution to the government budget is

$$B(a) = n^0(a)T(b) + [1 - n^0(a)] \int_0^1 T(w(a|p))dG(w(a|p)|a) - n^0(a)b - \bar{b}, \tag{36}$$

where $\bar{b}$ is a lump-sum transfer, see (1). The government budget is therefore $B = \int_0^1 B(a)dH(a)$.

Ideally, estimation of the model's structural parameters should be conducted under a balanced budget restriction, i.e. $B = 0$, achieved through the lump-sum transfers $\bar{b}$. However, these transfers transfers are not straightforward to implement in our model because a) we would need to take a stance on who who receives the lump-sum transfers, which in the absence of a government welfare function is somewhat arbitrary, and, more importantly, b) because recipients of government transfers will adjust their search effort as the income effects kick in. This, in turn, sets in motion a sequence equilibrium adjustments, which feeds back into the government budget. We are currently augmenting our equilibrium computation algorithm to allow for lump-sum transfers that balance the government budget. In doing so, we distribute any government surplus or deficit equally across workers, independently of ability. The current estimates, however, does not allow for lump-sum transfers, and hence, does not balance the government budget. That is, the current estimates impose $\bar{b} = 0$. To nonetheless facilitate interpretation, we always report changes in the government budget associated with various actual or counterfactual income tax reforms.

5.3 Auxiliary statistics

The estimation panel contains primarily information on wages and labor market transitions, and both these features of the data are needed to identify the model. In addition, the data has information on value added, which we also make use of. The description of the auxiliary statistics provided here is nontechnical and focuses on the intuition underlying the identification arguments for each of the structural parameters. As mentioned above, we do not currently attempt to identify the utility function (31).

Given our focus on a labor market with two-sided heterogeneity, it is useful to represent wages in a way that captures this feature of the data.[31] For this reason, let i index individuals, j index firms, and let n index annual November 28th cross sections. In addition, let $\mathrm{J}(i, n) = j$ if worker i in cross section n is employed by firm j, and consider the following auxiliary two-way fixed effect log-wage regression based on Abowd, Kramarz, and Margolis (1999),

$$\ln w_{in} = \phi_i + \psi_{\mathrm{J}(i,n)} + \varepsilon_{in}, \tag{37}$$

where, here, ϕ_i is a person effect, ψ_j is a firm effect, and ε_{in} represent residual log-wage variation. As indicated by the notation, we estimate this auxiliary model of wages using repeated annual

[31] Unlike wages, value added is available only at the firm-level, not at the match-level. Given the focus on two-sided heterogeneity, this limits the identifying power from value added data.

November 28th cross sections. Estimation of the worker and firm fixed effects is done by ordinary least squares imposing the assumption $\mathrm{E}[\varepsilon_{in}|i, \mathrm{J}(i,n)] = 0$, and we ensure that the design matrix of regressors, containing only worker and firm dummies, has full rank by estimating (37) only on the largest set of connected workers and firms, see Abowd, Creecy, and Kramarz (2002) for details.[32]

The auxiliary wage regression (37) allows a decomposition of log wages into a worker-component $\hat{\phi}_i$, a firm-component $\hat{\psi}_j$, and a residual component $\hat{\varepsilon}_{in} \equiv \ln w_{in} - \hat{\phi}_i - \hat{\psi}_{\mathrm{J}(i,n)}$. Indeed, $\ln w_{in} = \hat{\phi}_i + \hat{\psi}_{\mathrm{J}(i,n)} + \hat{\varepsilon}_{in}$. These components form the basis of many of the auxiliary wage statistics we include in the estimation. We stress that, even though (37) is not derived from the structural model and therefore likely is misspecified, the resulting decomposition of wages is useful because it yields measurements related to the underlying heterogeneity worker- and firm types, a and p.[33] We normalize the grand mean of the estimated firm effects to zero.

Identification of the match-level output function $y(a,p)$. The model-predicted wage offers $w(a|p)$ are complicated equilibrium objects that depends on all parameters of the model, including the match-level output function $y(a,p)$. However, holding search effort and hiring intensity fixed—and we argue below that search and hiring behavior is identified from data sources unrelated to wages—the distribution of wage offers is largely determined by the characteristics of the output function $y(a,p)$. Hence, statistics describing the distribution of wages contains identifying information regarding the parameters $\boldsymbol{\varrho}$ in the polynomial specification for $y(a,p)$, see (30). In addition to wage data, we make (at the moment, limited) use of information on value added per worker for the firms in our data to help identify $y(a,p)$.

Specifically, we include the average cross sectional wage, the 20th and 80th percentiles in the distribution of estimated worker effects from (37), and the same percentiles in the distribution of estimated firm effects. The average log wage in our data is 5.2273, and the 20th and 80th percentiles from the distribution of worker and firm fixed effects are -0.2754 and 0.2418 for the worker effects, and -0.0317 and 0.1631 for the firm effects. Hence, the wage data exhibits economically meaningful heterogeneity in both the worker and the firm dimension. Recall that the grand mean of the firm effects are normalized to zero, and that the reported percentiles therefore suggests the distribution of firm fixed effects has a rather long right tail.

We also include the so-called co-worker correlation. This latter statistics is the correlation between a individual's worker fixed effect and the average worker fixed effects of all his co-workers, see Lopes de Melo (2018). In our data, the co-worker correlation turns out to be 0.4421, suggesting a relatively strong tendency for high (low) worker fixed effect individuals to work with other high (low) worker fixed effect individuals. In our model, such sorting patterns are

[32] In our data, as well as in our model simulations, almost all workers and firms are connected.

[33] Indeed, our structural model implies that wages are increasing in firm-types (see Lemma 1), and simulations confirm that the model delivers wages that are also increasing in worker-types. Simulation studies in Bagger and Lentz (2015) indicate that the estimated "wage-types" from regressions like (37) captures "productivity-types" well provided wages are monotone in productivity-types.

primarily determined by differences in search behavior across worker-types, and these difference arise because the wage gains from reallocating from low- to high-type firms vary across the ability distribution, i.e. because of the modularity of the wage function $w(a|p)$. Since the wage offer distribution is endogenous, modularity of $w(a|p)$ reflect, in part, the modularity of the match-level output function.[34]

Regarding value added data, we include the average firm-level labor-share, i.e. the ratio of a firm's annual wage bill to it's annual value added, as well as the employment-weighted correlation between value added per worker-hour and the average wage paid by the firm. As detailed in the description of our in section 3, we observe total annual value added from a firm's VAT account with the Danish tax authorities, and the number of hours per year per firm is then computed using the hours information in the spell data *before* the sample selection described in section 3 takes place. The average labor share is 0.8028 in our data, and the employment-weighted correlation between average value added per worker and the average wage paid by the firm is 0.1360. We note that the correlation between observed firm productivity, i.e. value added per hour, and the average wage paid is positive, but not particularly strong.

Identification of the level of free search effort $\underline{s}$. The free search parameter $\underline{s}$ ensures all workers provide at least search effort $\underline{s}$, which first of all implies that the Diamond paradox is moot in the model. However, it also ensures that job search persists all the way to the top of the job ladder, which sustains an element of competition for workers at the higher rungs of the wage ladders. The competition manifest itself in the wage policies of firms, specifically in the variability of the firm fixed effects estimated from (37). The auxiliary statistics listed in the previous paragraphs relating to identification of $y(a,p)$ are numerous enough to over-identify $y(a,p)$, and we use these degrees of freedom to argue that the difference between the 20th and 80th percentile in the firm fixed effect distribution identifies $\underline{s}$. Simulation studies and our estimation procedure confirms this logic.

Identification of the disutility of job search effort function ζ_0 and ζ_1. The disutility from job search effort function $\zeta(\cdot)$ is central to our analysis and has two parameters, ζ_0 and ζ_1. ζ_0 determines the level of disutility, whereas ζ_1 governs the curvature, and therefore the substitution effect. If job search effort yields higher disutility at all effort-levels, job search effort will fall, both off- and on-the-job. Hence, higher disutility from job search, i.e. higher ζ_0, reduces the outflow from nonemployment. For a given job destruction rate, this increases the steady state nonemployment rate, see (14). Hence, we take identification of ζ_0 from the inclusion of the average nonemployment rate in our five-year panel, which is 0.1180.[35]

[34] Simulations confirm this to be the case, although we note that the Burdett-Mortensen wage setting game we apply here seems unable to deliver a submodular wage function, even when the output function is (strongly) submodular.

[35] The nonemployment rate is computed among workers who, at some point during our estimation period, are observed working. Our data contains individuals who are never working, so the unconditional nonemployment rate exceeds 0.1180. Our model interprets workers who are never, or rarely, working as being of types with very

Since ζ_1 governs the curvature of $\zeta(\cdot)$, it impacts differences in search effort between workers with different circumstances. Suppose the disutility function $\zeta(\cdot)$ is extremely convex. In that case, workers with very different returns to search will choose, essentially, the same job search effort—because the the disutility function is effectively "kinked". Hence, all else equal, an extremely convex disutility of search effort function, minimizes the variance of job durations, with all variation driven by variation in acceptance probabilities arising from wage dispersion. With less convex disutility, workers with different returns to search due to differences in ability and wages will make different choices regarding the search effort they provide. This increases the variance of job search effort, and therefore, all else equal, the variance of observed job durations. Based on that logic, we include the variance of the duration of the first job spell after non-employment in the vector of auxiliary statistics in order to identify ζ_1. In our estimation panel, where job durations are necessarily censored after five years of duration at the latest, we measure this variance to be 1.5107, including both censored and non-censored jobs durations.[36]

Identification of the hiring intensity cost function χ_1. Recall that the hiring cost function $\chi(\cdot)$ is normalized such that $\chi(1) = 1$, which leaves one parameter to be identified, namely χ_1, determining the curvature of the hiring cost function. This parameter is identified from the employment-weighted variance of firms' hiring rates. The logic is analogous to that applied above to argue identification of the curvature of the disutility of job search effort. We measure a firm's year t hiring rate as the ratio of new hires from November 28th, year t to November 27th, year $t+1$ to the number of employees in the firm on November 28th in year t, which allow us to compute firm-level hiring rates for 1999-2002.[37] The empirical employment-weighted variance of the hiring rate is 0.0827.

Identification of the sunk entry cost C. A high entry cost deters firm entry, which implies that the firms that do enter grow large. Hence, there is a positive relationship between the sunk entry cost parameter C and the average firm size in the economy. We consequently include the average firm size as an auxiliary statistic that we require our structural model to fit. A firm's size in year t is measured as the number of employees observed in the estimation data on November 28th in year t. The average firm size refers to the average taken across all firms and the five cross sections in our estimation panel ($t = 1999, 2000, ..., 2003$). We find an average firm size of 8.8414.

Identification of the job destruction process $\delta(a)$. Recall from (34) that we allow the job destruction process to be a function of worker ability. High-ability workers are over-represented high-wage workers, and we therefore use the empirical job destruction hazard conditional on

little return to job search.

[36]It is important to keep in mind that we replicate the data structure, and therefore this sampling of job durations, in the simulated data.

[37]We cannot compute the hiring rate for 2003, the final year in our estimation panel, because the 1999-2003 tax regime ends in December 31st, 2003.

workers' wages being below or above the median wage in the cross section. Specifically, consider all ongoing jobs at each annual cross section dates (November 28th), and compute the median wage in these jobs. Next, split the cross section sampled job spells according to whether the individual wage is above or below the median wage just computed. Finally, record the share of these jobs that end within a year because the worker transits to unemployment, for both below- and above median wage jobs. We find that the job destruction rate among jobs where workers are paid below the median wage is 0.2137 per year, whereas it drops to 0.0938 for jobs where workers are paid above the median wage. The model interprets the difference in job destruction rates as stemming from differences in the job destruction rate by worker ability.

Identification of the job reallocation rate μ. It is straightforward to show that, in a "standard" model of on-the-job search where job offers arrive at exogenous rates to nonemployed and employed workers, such as e.g. Burdett and Mortensen (1998), the share of jobs in a cross section of workers that end with the worker making a transition to nonemployment, is bounded from below by 1/2. With reallocation shocks this share may go below 1/2 because reallocations increases the prevalence of job-to-job transitions without increasing to the steady state mass of high-wage workers with high reservation wages. Hence, in this case, the reallocation rate is identified from the share of cross section jobs that end in a job-to-nonemployment transition—at least insofar as the empirical share falls below 1/2.

We have not established a lower bound for the steady state share of jobs that end in a job-to-nonemployment transition in our more complicated model with worker and firm heterogeneity, and endogenous search effort choices on both sides of the market, but our simulation studies suggest that the model cannot easily deliver a share below 1/2 without the additional job-to-job transition from reallocation shocks. Empirically, the share is 0.4745, i.e. close to, but below 1/2, thus providing identification of the reallocation rate μ.

5.4 Model fit

Having established identification of the model's structural parameters, we turn to estimation. Before reporting the structural parameter values, we describe the estimated model's fit to the auxiliary statistics just described. Table 6 reports the empirical and the simulated values for the auxiliary statistics used for estimation. In addition, Table 6 reports the diagonal elements in the Indirect Inference estimator's weight-matrix $\mathbf{\Omega}$. All off-diagonal elements in $\mathbf{\Omega}$ are set to zero. We put relatively high weight on average log wages and on the average firm size, and relatively little weight on the covariance between wages and hourly value added.

Overall, the fit reported in Table 6 is good. The over-weighting of average log wages ensures that the model accurately reproduces the level of wages. This is important as we applying the empirical tax function $T(\cdot)$ to the simulated wages. The fit to the 20th and 80th percentile in the distribution of worker fixed effects is spot on as well. However, while the model accurately reproduces the 20th percentile in the distribution of firm fixed effects, it does not generate the long right tail observed in the distribution of firm fixed effects estimated using the real data. The estimated structural model therefore underestimates the variability in log wages that can be attributed to firm fixed effects.

Furthermore, the model simultaneously underestimates the correlation between an individual's own worker fixed effect and her co-workers average worker fixed effect, and overestimates the correlation between the average wage paid by firms, and their value added. Bagger and Lentz (2015), employing a random search model bearing similarities to the the present model, also underestimated the "worker, co-worker" fixed effect correlation to the same extent, and attributes this to workers sorting on other dimensions than wages (or productivity). We believe the overestimation of the correlation between wages and value added largely stem from the absence of measurement errors in the simulated wage and output measurements. For this reason, we put relatively little weight on the "firm wage, value added"-correlation in the estimation.

With respect to the fit of the auxiliary statistics that are particularly informative about the disutility of job search effort function $\zeta(\cdot)$, namely the nonemployment rate and the variance of durations of jobs initiated by a nonemployment-to-job transition, we see that the estimated model does an excellent job in reproducing the empirical magnitudes. This is reassuring given that the curvature of the disutility job search effort is the key determinant of the substitution effect in job search effort in response to income tax reforms. The fit to the employment-weighted variance of the hiring rate, on the other hand, leaves room for improvement. The estimated model under-predicts the variance of hiring rates by more than 40 percent.

The average firm size carries relatively high weight in the estimation, and as a consequence, the fit is excellent, as is the fit to the wage conditional job destruction hazards, even though these latter two moments carries normal weight in the estimation. Finally, the share of jobs ending in job-to-nonemployment transition, an auxiliary statistic included to identify the reallocation rate μ, is fitted reasonably well, although the estimated model's prediction exceeds 1/2, while the data, as noted above, has a share slightly below 1/2.

5.5 Structural parameter estimates

A subset of the structural parameter estimates are reported in Table 7, along with their asymptotic standard errors (see footnote 29). We do not report the parameter estimates of δ_0 and δ_1 that shape the job destruction process $\delta(a)$, nor of $\boldsymbol{\varrho}$, the parameters of the match output function $y(a, p)$, both which we instead render graphically in Figure 3.

Turning attention first on the estimates reported in Table 7, we note that, with an estimate of ζ_1 of 1.5257, the elasticity of job search effort disutility with respect to job search effort is 1.66, i.e. with slightly less curvature than a quadratic disutility function. This is consistent with the empirical results in Christensen, Lentz, Mortensen, Neumann, and Werwatz (2005), Kreiner, Munch, and Whitta-Jacobsen (2015), and Bagger and Lentz (2015), who all utilizes Danish data, and who all report approximately quadratic job search cost functions. We also note that ζ_0 is precisely estimated at 0.1011, that free search $\underline{s}$ is estimated to be only 0.0212,

Table 6: Model fit

	Simulated	Data	Weight
Average log wage	5.2254	5.2273	100.00
Workers' share of value added	0.8407	0.8028	1.00
20th percentile, worker fixed effects	−0.2795	−0.2754	1.00
80th percentile, worker fixed effects	0.2405	0.2418	1.00
20th percentile, firm fixed effects	−0.0339	−0.0317	1.00
80th percentile, firm fixed effects	0.0606	0.1631	1.00
Corr(worker fixed effect, average co-worker fixed effects)	0.2681	0.4421	1.00
Corr(average firm wage, hourly value added)	0.2447	0.1360	0.10
Nonemployment rate	0.1289	0.1180	1.00
Variance, duration of the 1st job out of nonemployment	1.6097	1.5107	1.00
Variance, hiring rate (employment weighted)	0.0471	0.0827	1.00
Average firm size	8.9939	8.8414	10.00
Job destruction hazard if wage $<$ median wage	0.1799	0.2137	1.00
Job destruction hazard if wage $\geq$ median wage	0.0918	0.0938	1.00
Share of jobs ending in job-to-nonemployment transition	0.5517	0.4745	1.00

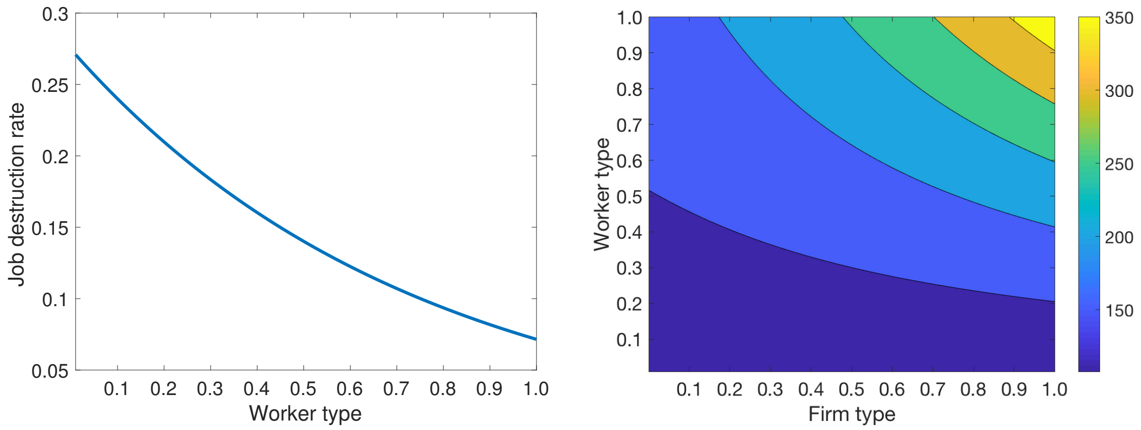
Note: The column labeled "weight" reports the diagonal elements in the weight-matrix $\mathbf{\Omega}$, see (35). Off-diagonal entries in $\mathbf{\Omega}$ are set to zero.

Table 7: Structural parameter estimates

Parameter	Estimate
Disutility of job search effort $\zeta(s - \underline{s}) = \frac{\zeta_0}{1+1/\zeta_1} s^{1+1/\zeta_1}$	
Level, ζ_0	0.1011 (0.0191)
Curvature, ζ_1	1.5257 (0.2205)
Free search, $\underline{s}$	0.0212 (0.0010)
Reallocation rate (annual), μ	0.1489 (0.0339)
Hiring intensity cost $\chi(v) = \frac{1}{1+1/\chi_1} s^{1+1/\chi_1}$	
Curvature, χ_1	0.7244 (0.0390)
Sunk entry cost, C	$5,082$ (10.4026)

Note: Asymptotic standard errors are reported in parentheses.

Figure 3: Job destructions and match output



Panel A: Annual job destruction rate $\delta(a)$ Panel B: Match output function $y(a, p)$

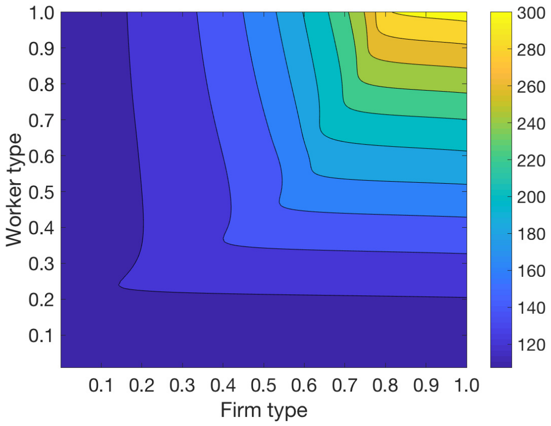
Note: Match output is denominated in Danish Kroner per hour.

and that the job reallocation rates μ is estimated to be 0.1489. The latter three parameters, however, can be interpreted only in proportion to the baseline job offer arrival rate $\lambda(a)$, an equilibrium object that we return to below. The final two parameter estimates reported in Table 7 refer to the firm-side of the model. Hiring cost is convex with the elasticity with respect to hiring effort being 2.38 and the sunk entry cost is estimated to be $5,082$ Danish Kroner.

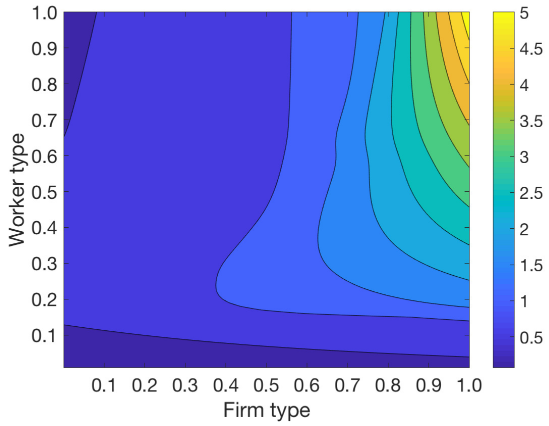
Panel A in Figure 3 plots the estimated annual job destruction process as a function of worker ability a. The job destruction rate exhibits a strong negative relationship with worker ability a. Indeed, a worker at the 25th percentile in the ability distribution has her job destroyed, on average, every fifth year, whereas a worker at the 75th percentile waits, on average, 10 years between job destruction events. Panel B in Figure 3 is a contour plot of the estimated match output function $y(a, p)$. By assumption, $y(a, p)$ is increasing in both arguments, but it is evident that the estimated match output function is supermodular: The increase in output from reallocating a high-ability worker from a low- to a high-productive employer is greater than the increase realized by the same reallocation of a low-ability worker. Likewise, the increase in output from replacing a low-ability worker with a high-ability worker is greater for firms with higher productivity. This completes the description of the model's exogenous components.

The estimated model favors high-type workers in two important ways. First, high-type workers face lower job destruction rates than low-type workers. All else equal, this ensures that, on average, high-type workers are matched with better firm-types than low-type workers. It also increases the return to job search for high-type workers vis-a-vis low-type workers. Second, the match output function is increasing in a by construction, and in addition, is estimated to be supermodular. Insofar as the supermodularity is carried over to the equilibrium wage function, and provided income effects, and any countervailing equilibrium effects running through firms hiring intensity choices $v(a|p)$ or the equilibrium baseline job offer arrival rate $\lambda(a)$ are relatively small, this will tend to further widen the gap in returns to job search effort between high- and

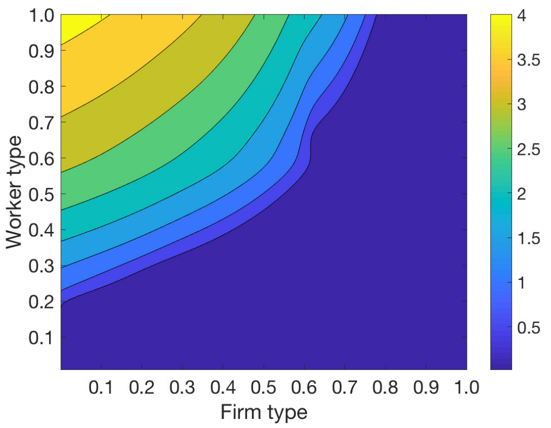
Figure 4: Key equilibrium objects



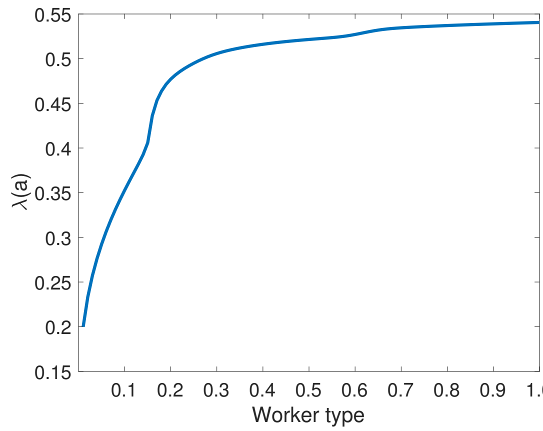
Panel A: Wage offers $w(a|p)$



Panel B: Vacancy intensity $v(a|p)$



Panel C: Job search effort $s(p|a)$



Panel D: Baseline job finding rate $\lambda(a)$

Note: Wages are denominated in Danish Kroner per hour.

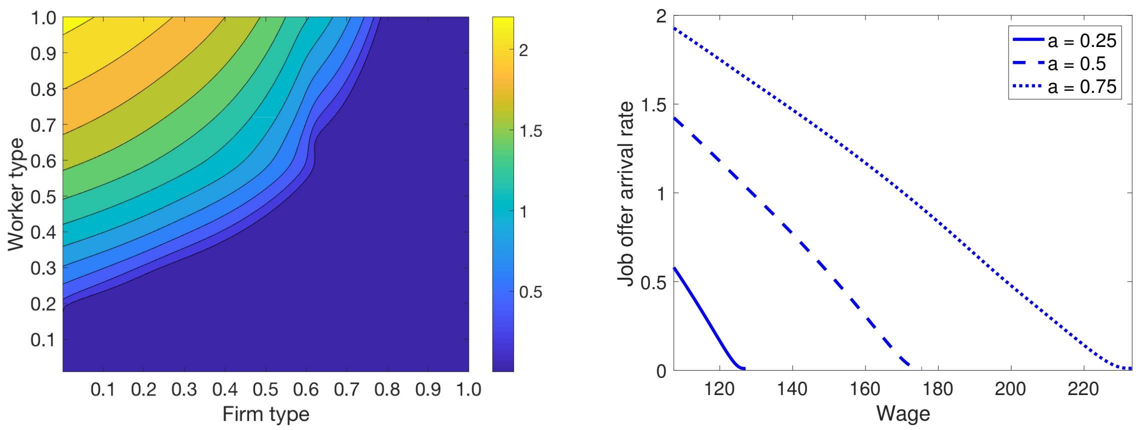
low-type workers.

Figure 4 depicts a set of key equilibrium objects relevant for understanding the implications of the estimated model for the equilibrium allocation of labor. Panel A plot the contours of the estimated equilibrium wage function $w(a|p)$. Of course, as established in Lemma 1, wages are strictly increasing in p, but, in addition, higher-type workers almost always reap greater wage gains by moving from a low- to a high-productive firm than do workers of lower ability. Unlike the estimated match output function, the wage function is, however, not globally supermodular.[38] Panel B plots a contour plot of firms' equilibrium hiring intensity as a function of worker- and firm-types. Consistent with Lemma 2, hiring intensity $v(a|p)$ is always increasing in p for a given a, but we see that hiring intensity $v(a|p)$ need not be increasing in a for a given p, although there is a very strong tendency for this to be the case.

Turning to the key decision margin of workers—job search effort—Panel C in Figure 4

[38] We are not entirely sure what causes this feature of the equilibrium wage function, and cannot at the time of writing rule out that it is caused by numerical problem in computing the equilibrium wage function.

Figure 5: Estimated annual job offer arrival rates by worker- and firm-type



Panel A: $\lambda(a)s(p|a)$ Panel B: $\lambda(a)s(w|a)$

Note: Panel A plots the estimated job offer arrival rates as a function of firm-type $p \in [p_0, 1]$ and the Panel B plots the estimated job offer arrival rates as a function of wage $w = w(a|p) \in [b, \overline{w}(a)]$.

contains a contour plot of workers' optimal job search effort as a function of worker- and firm-types. As stipulated by Lemma 1, job search effort is strictly decreasing in p for each a with optimal search effort in the most productive firm being $\underline{s}$ for all a. Moreover, we see that, at any firm-type p, higher-type workers exert more job search effort than lower-type workers. Hence, the equilibrium implies higher return to search for high-type workers.

As noted further above, the magnitude of job search effort, free search, reallocation shocks must be interpreted in proportion to the relevant baseline job offer arrival rate $\lambda(a)$. Panel D in Figure 4 plots $\lambda(a)$ as a function of a. We see that $\lambda(a)$ is strictly increasing and concave function in a. The labor market for the least able workers receive job offers at an annual rate of 0.2 per unit of search, whereas, at the other end of the ability spectrum, high-ability workers receive offers at annual rate 0.55 per unit of search, emphasizing again the estimated equilibrium is such that return to job search effort is higher for high ability workers.

Finally, Figure 5 puts Panels C and D from Figure 4 together and plots the annual job offer arrival rates resulting from on-the-job search effort implied by the estimated model. Panel A plots the job offer arrival rate as a function of firm-type, $\lambda(a)s(p|a)$ by way of a contour plot. Because both $s(p|a)$ and $\lambda(a)$ are increasing functions of a, conditional on the type of the employing firm, workers of higher ability will, in equilibrium, receive more job offers than workers of lower ability. Panel B in Figure 5 plot the job offer arrival rate for workers of type $a = 0.25$, $a = 0.50$ and $a = 0.75$, respectively, as a function of the wage, i.e. $\lambda(a)s(w|a)$ where $w = w(a|p)$. We see that, at any wage rate, more able workers receive more job offers. Indeed, at the lowest wage, equilibrium behavior implies that a type $a = 0.75$ worker is four times more likely to receive a job offer than a type $a = 0.25$ worker. We stress that these differences arise endogenously, and that they have implications for the equilibrium allocation of labor, a subject we turn to next.

5.6 The equilibrium allocation of labor

The allocation of labor is represented by the worker-ability specific nonemployment rates, $n^0(a)$ given by (14), and the worker-ability specific distributions of workers across wage rates w, $G(w|p)$ given by (16). Note that, since $w = w(a|p)$ is a one-to-one function of p, the distribution of worker-types across firm types is simply given as $G(p|a) = G(w(a|p)|a)$.

We have already seen from Table 6 that the estimated model implies an empirically reasonable aggregate nonemployment rate $n^0 = \int_0^1 n^0(a)dH(a)$ around 0.13 among those workers ever observed in employment over the 5 years of simulated data. However, the aggregate nonemployment rate masks heterogeneity across worker-types. Panel A in Figure 6 plots the estimated steady state nonemployment rates as a function of worker ability a. We note that the estimated model implies that the workers in the lowest quintile of the ability distribution face very high nonemployment rates, in excess of 20 percent, while workers in the top quintile of the ability distribution has nonemployment rates around 2 percent.[39]

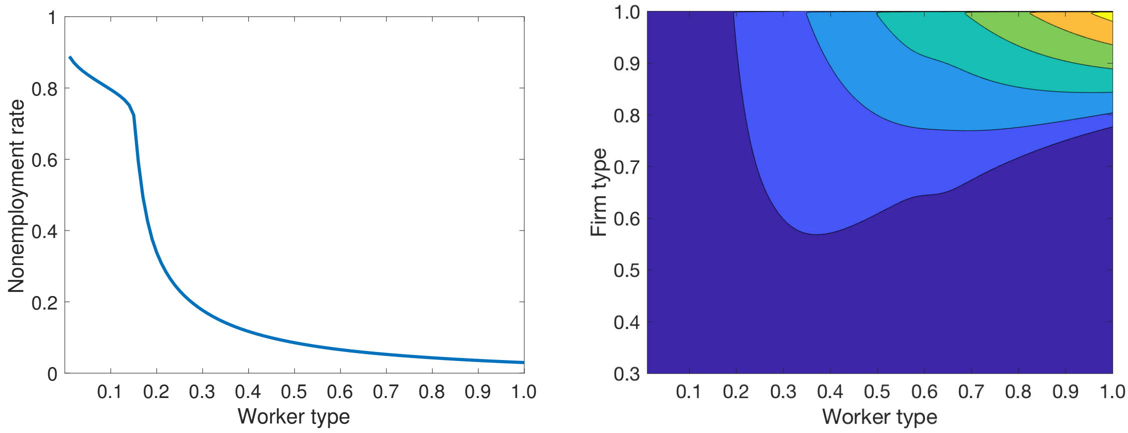
Panel B in Figure 6 illustrates the conditional densities $g(p|a)$ as a function of a. We see a clear pattern whereby high-type workers tend to cluster in high-type firms to a much higher extent than low-type workers. In that sense, the equilibrium allocation exhibits traits of positive assortative matching, a finding that is consistent with Bagger and Lentz (2015). This equilibrium sorting results come about because high ability workers have higher returns to job search effort, due to their lower unemployment risk, see Panel A in Figure 3, and the supermodularity of the match output function, see Panel B in Figure 3. The endogenous response to the higher returns to job search effort result in higher job offer arrival rates, see Panel C in Figure 5, which implies that high-ability workers climb the wage (or firm productivity) distribution faster than low-ability workers. The remainder of the paper analyzes and quantifies the impact of the income tax schedule on the equilibrium allocation of labor.

6 Equilibrium tax reform evaluations

We initiate our analysis of the impact of labor income taxation on the allocation of labor by employing the estimated structural model to conduct equilibrium evaluations of a series of Danish income tax reforms during 1990-2005. The reforms under consideration took place in 1990, 1994, and 2004 and the context in which they were devised and implemented is described in section 4.1. Figure 7 shows the marginal tax functions (Panel A) and the average tax functions (Panel B) for the four tax regimes we consider, including the 1999-2003 regime used for estimation. Inspection of Figure 7 reveals that each of the implemented changes in the income tax system during 1990-2005 (almost) uniformly flattened the tax function, i.e. lowered the marginal rate at all wage rates. At the same time, the average tax rate was increased at

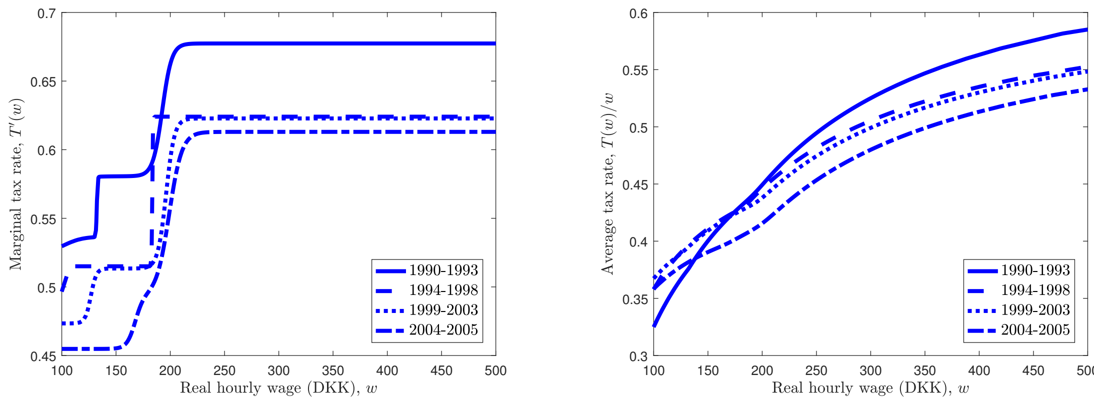
[39] Workers at the very bottom of the ability distribution are effectively estimated to be nonparticipants. Many of these workers will not be observed in employment during a 5-years simulation window, and will therefore not contribute to the nonemployment rate of 0.13 reported in Table 6. Our data, of course, also contains nonparticipants.

Figure 6: The equilibrium allocation of labor



Panel A: $n^0(a)$ Panel B: $g(p|a)$

Figure 7: Danish tax regimes 1990-2005: Marginal and average tax functions



Panel A: $T'(w)$ Panel B: $T(w)/w$

Note: Panel A plots the marginal tax function $T'(w)$ and Panel B plots the average tax function $T(w)$.

the bottom of the wage distribution, and lowered everywhere else.

The equilibrium evaluations are conducted as comparative statics between the steady state equilibria associated with each of the income tax regimes. The structural parameters are held fixed throughout. Our interest centers on aggregate and worker-type specific steady state, i.e. long run, nonemployment rates and expected labor income. The worker-type specific nonemployment rate $n^0(a)$ is given by (14), with the aggregate nonemployment rate given by $\int_0^1 n^0(a)dH(a)$. The worker-type a specific expected labor income is given by

$$LI(a) = [1 - n^0(a)] \int_0^1 w(a|p)dG(w(a|p)|a), \tag{38}$$

where $G(w|a)$ is given by (16), and aggregate labor income is $LI = \int_0^1 LI(a)dH(a)$.

The tax reforms also impact the government budget. As mentioned above, our current equilibrium computation algorithm does not feature lump-sum transfers to balanc the government budget. Instead, to at least provide context for the reported effects on the equilibrium labor allocation, we also report the change in government budget associated with each of the tax regime comparisons.

Table 8 reports the percentage change in nonemployment n^0, labor income LI, and the government budget B between each of the 1990, 1994, and 2004 Danish income tax regimes, and the baseline 1999 regime. Hence, in Table 8, the 1999 tax regime is always the baseline. This means that the numbers reported in the 1990-1993 panel evaluates the effect of changing the 1999-2003 regime (back) to the 1990-1993 regime, while the numbers in the 2004-2005 panel evaluates the effect of the actual reform 2004 reform, i.e. changing the 1999-2003 regime to the 2004-2005 regime. Perhaps confusingly, a negative number for, say, labor income in the 1990-1993 panel therefore means that the allocation *improved* between the 1990-1993 regime and the 1999-2003 regime, while a negative number for labor income in the 2004-2005 panel means that the allocation of labor *deteriorated* between the 1999-2003 regime and the 2004-2005 regime.

We consider the aggregate response (rows labeled "All" in Table 8), as well as the responses at the 25th, 50th, and 75th percentile of the ability distribution (rows labeled 0.25, 0.50, and 0.75, respectively). For each variable, we consider the total effect, comprising both the substitution effect and the income effect, as well as the substitution effect in isolation. Hence, the difference between pairs of columns labeled "Total" and "Subst" in Table 8 is the income effect.[40] Finally, we consider both the partial equilibrium effect (labeled "PE") and full equilibrium effect (labeled "FE"). The partial equilibrium effect holds firm behavior and equilibrium variables constant, allowing only workers to respond to changes in the income tax code.

Long run nonemployment. Relative to the 1999-2003 tax regime, the 1990-1993 tax regime involves a 4.9 percent higher steady state nonemployment rate when accounting for equilibrium effects. The corresponding partial equilibrium effect is much larger at 7.4 percent. Income effects are modest at the aggregate level. In partial equilibrium, nonemployment increases 7.9 percent due to the substitution effect. This causes an increase in the marginal value of consumption, and the resulting income effect in workers' job search effort reduces the total effect effect slightly to 7.4 percent. The full equilibrium adjustment is such that the income effect in the aggregate, albeit small, works in the same direction as the substitution effect. If we look at the effect within the ability distribution, we see that workers at the bottom of the ability distribution experience the largest increase in nonemployment. The differences across ability levels is especially pronounced in full equilibrium. We also note that the positive aggregate income effect in full equilibrium is driven by equilibrium adjustments at the bottom of the ability distribution—towards the top of the ability distribution, the income effect is negative, dampening the substitution effect.

Similar results, although with smaller magnitudes are found when comparing the 1994-1998 regime to the baseline 1999-2003 regime. Although full equilibrium income effects are still

[40]The substitution effect holds the marginal utility of consumption constant, see section 2.2.2.

Table 8: Evaluating the 1990, 1994, and 2004 Danish income tax reforms for nonemployment, labor income and the government budget

		$\frac{\Delta n^0}{n^0} \times 100\%$				$\frac{\Delta LI}{LI} \times 100\%$				$\frac{\Delta B}{B} \times 100\%$			
		Total		Subst		Total		Subst		Total		Subst	
Comparison tax regime	Ability	PE	FE	PE	FE	PE	FE	PE	FE	PE	FE	PE	FE
1990-1993	All	7.4	4.9	7.9	4.8	−1.0	−1.8	−1.0	−1.9	−1.6	−2.7	−1.5	−2.5
	0.25	10.3	9.9	7.8	5.7	−3.1	−3.9	−2.4	−2.2	−22.9	−24.1	−20.2	−18.9
	0.50	9.3	5.3	9.3	5.4	−1.0	−1.7	−1.0	−1.8	−2.4	−2.8	−2.4	−2.8
	0.75	8.1	3.5	10.4	5.6	−0.6	−1.3	−0.8	−1.8	2.3	1.6	2.0	0.9
1994-1998	All	1.7	1.1	1.9	1.2	−0.3	−0.5	−0.3	0.5	−0.3	−0.6	−0.3	−0.6
	0.25	3.3	3.2	2.9	2.6	−1.0	−1.2	−0.9	1.0	−5.7	−6.1	−5.2	−5.3
	0.50	1.8	0.8	1.7	0.7	−0.2	−0.3	0.9	0.8	−0.6	−0.6	−0.6	−0.5
	0.75	1.8	0.6	2.3	1.1	−0.1	−0.4	−0.2	−0.5	0.7	0.5	0.6	0.3
2004-2005	All	−1.7	−1.4	−3.6	−2.6	0.2	0.3	0.4	0.7	−5.1	−5.0	−4.6	−4.4
	0.25	−1.4	−1.7	−2.6	−2.4	0.4	0.8	0.8	1.1	−4.0	−3.2	−2.7	−2.4
	0.50	−2.9	−2.3	−5.4	−4.0	0.3	0.5	0.6	1.0	−4.8	−4.7	−4.3	−4.1
	0.75	−1.9	−1.3	−4.5	−2.8	0.1	0.0	0.3	0.5	−4.4	−4.5	−4.1	−4.0

Note: The percentage changes are relative to the 1999-2003 income tax regime. Hence, positive (negative) numbers in the 2004-2005 panel indicates that the variable increased (decreased) going from the 1999-2003 regime to the 2004-2005 tax regime. Positive (negative) numbers in the 1994-1998 and 1990-1993 panels, however, indicate that the the variable decreased (increased) as a result of going from the 1994-1998 (or 1990-1993) tax regime to the 1999-2003 regime.

positive towards the bottom of the ability distribution, the aggregate income effect is negative, but still small. In full equilibrium, aggregate nonemployment is 1.1 percent higher in the 1994-1998 regime than in the 1999-2003 regime. The distinction between partial and full equilibrium adjustments continues to be quantitatively important. In comparison to the 1999-2003, the 2004-2005 regime involves 1.4 percent lower nonemplyment in full equilibrium, with the full equilibrium substitution effect being a 2.6 percent lower nonemployment rate. In the case of the 2004-2005 regime, income effects become economically significant, often reducing the substitution effect by up the 50 percent.

Overall, each of the 1990, 1994, and 2004 Danish income tax reforms lead to a reduction in nonemployment. The effect reported in Table 8 implies that the tax reforms implemented during 1990-2005 in Denmark reduced steady state nonemployment by approximately 6 percent. However, towards the bottom of the ability distribution, the reduction was larger, around 13 percent at the 25th percentile in the ability distribution, while the effect is smaller for workers with higher ability, around 3% among workers at the 75th percentile in the ability distribution. The income effect reduces the effect for high ability workers in particular.

Long run labor income. Consider first the evaluation of the 1990-1993 tax regime relative to the baseline 1999-2003 regime in the first panel of Table 8. Table 8 shows that the overall effect amounts to a 1.8 percent drop in labor income, accounting for equilibrium adjustments. The partial equilibrium response is a 1 percent drop in labor income, again highlighting the importance of modeling the equilibrium response to tax reforms. On the aggregate, the income effects are small. However, we again find heterogenous effects across the ability distribution with low ability workers experiencing a much larger reduction in their labor income than workers of higher ability.

Turning now to the comparison between the 1994-1998 regime relative to the baseline 1999-2003 regime, we obtain a smaller increase in aggregate labor income of 0.5 percent accounting for equilibrium adjustment. The partial equilibrium response indicates a 0.3 percent decrease in labor income. Accounting for the income effects in the comparison between the 1994-1998 regime and the 1999-2003 regime is quantitatively important. The substitution effect indicate a 0.5 increase in labor income, implying that the income effect reverses the sign of the predicted labor income response. The labor income response again exhibits heterogeneity across the ability distribution.

Finally, consider the comparison of labor incomes between the 2004-2005 tax regime and the baseline 1999-2003 regime, i.e. the third panel in Table 8. Going from the 1999-2003 tax regime to the 2004-2005 regime involves a 0.3 percent long run gain in aggregate labor income, when equilibrium adjustments on the labor market are taken into account. These equilibrium adjustments are, however, small, with the partial equilibrium effect being a 0.2 percent increase. The relatively small total effect masks a large substitution effect increasing steady state labor income by 0.7 percent. Again, we see evidence of heterogenous responses throughout the ability distribution.

The overall long run effect on labor income of the tax reforms implemented in Denmark between 1990 and 2005 is an approximate 2 percent increase, with larger effects (an approximate 4.5 percent long run increase) at the 25th percentile of the ability distribution, whereas for the more able workers at the top of the ability distribution, smaller effects materialize (an approximate 1 percent increase).

Long run government budget. The third set of columns in Table 8 present the percentage change in the government budget, comparing the 1990-1993, the 1994-1998, and the 2004-2005 tax regimes to the baseline 1999-2003 regime. The reported figures imply that, in full equilibrium and accounting for income effects, the government budget increased 2.7 percent when going from the 1990-1993 regime to the 1999-2003 regime. If we consider the effect of going from the 1990-1993 to the 1999-2003 regime on the ability-specific contributions to the government budget, we see a large increase in tax revenue at the bottom of the ability distribution, with the effect tapering off as we move up through the distribution.

The budgetary implications of going from the 1994-1998 regime to the 1999-2003 regime are quite small, and we shall not comment further on them here, except to note that, accounting for equilibrium effects, the government budget increased 0.6 percent as a result of going from the 1994-1998 regime to the 1999-2003 regime. Keeping in mind that the average tax rate was reduced everywhere, except at quite low wage levels, it is interesting to note that government budget nonetheless increased as a result, perhaps suggesting the Danish labor market operated to the right of the peak of the Laffer-curve prior to 1999.[41]

Finally, consider the budgetary implications of going from the 1999-2003 regime to the 2004-2005 regime. From Table 8, we see that this reform leads to a rather large declines in government revenue, of 5 percent in full equilibrium, with the partial equilibrium effect, and the pure substitution effects being of the same magnitude. Considering Panel B in Figure 7, we note that the going from the 1999-2003 regime to the 2004-2005 regime involves uniformly reducing the average tax rates at all wage levels. The improved allocation is simply not enough to compensate for the associated mechanical loss. Given the sizeable budgetary implications involved in the 2004 income tax reform, one might speculate that allowing for lump sum transfers to balance the budget may impact the conclusions we reached for this reform regarding nonemployment and labor income.

Summing up. A series of income tax reforms in Denmark implemented in 1994, 1999, and 2004 improved the equilibrium allocation of labor, reducing unemployment by about 6 percent, and increasing labor income by about 2 percent, with workers of low ability gaining the most. Our analysis does not balance the government budget, and overall, the tax reforms during

[41]In our model, the "Laffer effect" is operates through labor allocation adjustments. By lowering taxes, the government sustains a mechanical loss in revenue, but the behavioral effects on workers' job search behavior improves the allocation of labor, leading to increases in government revenue. The behavioral gain may dominate the mechanical loss.

1990-2005, lead to a 2 percent drop in government revenue. Our analysis has shown that the key ingredients in our model—two-sided heterogeneity, taking account of possible income effects, and equilibrium adjustments—are quantitatively important for understanding the long run labor market adjustments to the implemented income tax reforms. Appendix E provides a complementing analysis of counterfactual linear income tax functions to assess the impact of progressivity on the equilibrium allocation of labor.

7 The long run elasticity of taxable labor income

The evaluation exercises in the previous section demonstrated that, when analyzed through the lens of our structural model, the long run allocation of labor respond in economically meaningful ways to the incentives provided by the income tax system. The measured response operates through the job search effort choices of workers, and the hiring intensity choices and wage polices of firms, or put succinctly, through labor mobility.

Following Feldstein (1995, 1999), the broader public economics literature typically summarizes economic responses to taxation in the elasticity of taxable income, or more recently the elasticity of taxable labor income. The shift in focus towards taxable labor income reflects that labor income is less prone to evasion and avoidance, which means that behavioral responses can be understood within standard models of the labor market, see e.g. Chetty, Friedman, Olsen, and Pistaferri (2011). From a theoretical point of view, the elasticity of taxable labor income comprises all behavioral responses to changes in the tax system of relevance for labor income, including hours responses, and effort provision in various dimensions. Empirically, Kleven and Schultz (2014) report estimates of the taxable labor income elasticity in Denmark ranging between 0.05 and 0.12.[42]

Estimates of the taxable labor income elasticity, such as the ones reported in Kleven and Schultz (2014), are often obtained using variation in marginal tax rates induced by tax reforms as a sources of exogenous variation in incentives across individuals. As discussed in Kreiner, Munch, and Whitta-Jacobsen (2015), such an approach is, however, unlikely to capture the search induced labor allocation effects. Indeed, Kreiner, Munch, and Whitta-Jacobsen (2015) uses a partial equilibrium job search model to estimate the long run elasticity of taxable labor income for Denmark, accounting for steady state labor allocation responses, at 0.30. The model used for estimation in Kreiner, Munch, and Whitta-Jacobsen (2015) does not feature equilibrium responses, does not explicitly account for worker heterogeneity, and does not allow for income effect in the job search effort responses of workers. Our model incorporates all these features. It is therefore of interest to revisit the elasticity of taxable labor income within the context of our richer model

To proceed, recall that, within our model, taxable labor income $LI = \int_0^1 L(a)dH(a)$, where $LI(a)$ is given by (38). $LI(a)$ depends on the equilibrium allocation of labor across wage levels

[42] For the US, the elasticity of taxable income is typically estimated to be around 0.25, see Saez, Slemrod, and Giertz (2012).

Table 9: Long run taxable labor income elasticities

	Total		Subst	
Ability	PE	FE	PE	FE
All	0.069	0.123	0.091	0.169
0.25	0.217	0.289	0.241	0.297
0.50	0.068	0.110	0.093	0.156
0.75	0.031	0.081	0.060	0.142

$G(w|a)$, see (16). The labor allocation, in turn, depends on workers' search effort which may be distorted by income taxation. Following Kreiner, Munch, and Whitta-Jacobsen (2015), we compute the elasticity of taxable labor income for type-a workers, by recording the percentage change in their labor income $LI(a)$ following a uniform 1 percent increase in the marginal net-of-tax rate $1 - T'(w)$, keeping $b - T(b)$ constant.[43] Let $\epsilon(a)$ denote the computed elasticity of taxable labor income for type-a workers. The aggregate elasticity of taxable labor income is then $\epsilon = \int_0^1 \epsilon(a) dH(a)$.

Table 9 reports the partial and full equilibrium elasticities implied by the estimated structural model, both aggregate, and for workers at the 25th, 50th and 75th percentile in the ability distribution. Table 9 reports both the total effect, comprising both a substitution and an income effect, as well as elasticity estimates based purely on the substitution effect.

The results in Table 9 echoes a common theme running through this paper, namely that the distinctions between partial and full equilibrium, income and substitution effects, as well as worker heterogeneity are quantitatively important. We estimate the full equilibrium mobility-based elasticity of taxable labor income, comprising both substitution and income effects, to be 0.123, a substantially lower elasticity estimate than that reported by Kreiner, Munch, and Whitta-Jacobsen (2015). When we shut down equilibrium adjustments and consider only the substitution effect to facilitate closer correspondence to the modeling framework employed in Kreiner, Munch, and Whitta-Jacobsen (2015), we obtain an elasticity of 0.091.[44] In full equilibrium, the uncompensated elasticity of taxable labor income is larger, at 0.169. Finally, from Table 9, we note a clear pattern, whereby workers of lower ability have higher taxable labor income elasticities. Indeed, workers at the 25th percentile in the ability distribution have taxable labor elasticities close to the 0.30 reported in Kreiner, Munch, and Whitta-Jacobsen (2015), whereas workers at the 75th percentile have three times lower elasticities, at 0.081.

[43] That is, we change the net-of-tax rate faced by workers from $1 - T'(\cdot)$ to $1.01 \times [1 - T'(\cdot)]$.

[44] It is at present not clear to us why the difference between the two estimates arises. We have estimated the Kreiner, Munch, and Whitta-Jacobsen (2015) model on our data and obtained results similar to those they report, ruling out differences in the data we use. Kreiner, Munch, and Whitta-Jacobsen (2015) do, however, use a different tax function than we do, but we have yet to ascertain that differences in the tax function are responsible for the different estimates.

7.1 Decomposing the long run elasticity of taxable labor income

Our rich structural model allow us to further decompose the elasticity of taxable income to gain further insights into the behavioral (mobility) responses to income taxation. For this purpose, let $LI(a)$ and $\hat{LI}(a)$ be expected taxable incomes of a type-a worker under two different income tax regimes, $T(\cdot)$ and $\hat{T}(\cdot)$. Also, let $n^0(a)$ and $\hat{n}^0(a)$, $G(p|a)$ and $\hat{G}(p|a)$, and $w(a|p)$ and $\hat{w}(a|p)$ be the nonemployment rates, wage distribution, and wage policies under the $T(\cdot)$ and $\hat{T}(\cdot)$ regimes, respectively. Given the expression for the expected taxable income of a type-a worker, (38), the following decomposition of the change in taxable labor income between the two tax regimes $T(\cdot)$ and $\hat{T}(\cdot)$ is of interest:

$$\begin{aligned} LI(a) - \hat{LI}(a) &= \underbrace{[n^0(a) - \hat{n}^0(a)] \int_0^1 \hat{w}(a|p)\, d\hat{G}(p|a)}_{\text{Extensive margin allocation}} \\ &+ \underbrace{[1 - n^0(a)] \int_0^1 \hat{w}(a|p)\, [\hat{g}(p|a) - g(p|a)]dp}_{\text{Intensive margin allocation}} \\ &+ \underbrace{[1 - n^0(a)] \int_0^1 [\hat{w}(a|p) - w(a|p)]\, dG(p|a).}_{\text{Wage policy margin}} \end{aligned} \tag{39}$$

Equation (39) states that the change in taxable labor income between the two tax regimes $T(\cdot)$ and $\hat{T}(\cdot)$ can be decomposed into responses along three different margins. First, a change in income taxation brings about a change in the equilibrium nonemployment rate, an extensive margin allocation effect. Second, a change in income taxation also shifts the equilibrium distribution of workers across firms, denoted the intensive margin allocation effect. Third, a change in income taxation induces changes in the equilibrium wages offered by firms, an effect we denote the wage policy margin.[45]

With (39) in hand, we can decompose the elasticity of taxable labor income estimates provided in Table 9. We focus here exclusively on the decomposition of the elasticity accounting for equilibrium adjustments, which was found to be 0.123 at the aggregate in Table 9.[46] The decomposition is reported in Table 10, where we see that 46 percent of the aggregate taxable labor income elasticity can be ascribed to reallocations on the extensive margin, whereas 12 percent can be ascribed to reallocations on the intensive margin. Adjustments on the wage policy margin account for the remaining 42 percent of the estimated elasticity of taxable labor income.

When we apply the decomposition (39) to taxable labor income elasticities across the ability distribution an interesting pattern emerge. Adjustments along the extensive margin are respon-

[45] The labor income decomposition (39) is carried out such that the extensive margin allocation is evaluated at the intensive allocation and the wage policy under $\hat{T}(\cdot)$, and the intensive allocation margin is evaluated at the extensive margin allocation under $T(\cdot)$, and the wage policy under $\hat{T}(\cdot)$. The wage policy margin is evaluated at the allocation under $T(\cdot)$. The decomposition is robust to the order in which we evaluate the three terms.

[46] Decompositions of the partial equilibrium elasticities yield qualitatively similar results and are available upon request.

Table 10: Decomposing the full equilibrium elasticity of taxable labor income

	All	$a = 0.25$	$a = 0.50$	$a = 0.75$
Elasticity of taxable labor income	0.123	0.289	0.110	0.081
Extensive margin allocation	46%	73%	32%	15%
Intensive margin allocation	12%	2%	15%	21%
Wage policy margin	42%	24%	53%	64%

sible for the bulk of the measured taxable labor income elasticity at the bottom of the ability distribution. In Table 10, extensive margin adjustments account for 73 percent of measured elasticity among workers at the 25th percentile of the ability distribution. This share declines monotonically to 15 percent once we consider workers at the 75th percentile of the ability distribution. By implication, the opposite pattern emerges when we consider adjustments along the intensive margin, i.e. changes in $G(p|a)$ and wage policies, i.e. changes in $w(a|p)$. The relative importance of both of these two margins of adjustment increases monotonically in the ability of workers.

8 A Pareto optimal tax reform

We have throughout the paper refrained from discussing normative issues, i.e. optimal income taxation. Our model has an exclusively positive focus and is designed to capture the enormous amount heterogeneity present in labor market data, and to deliver empirically realistic behavior, including wage distributions. As result, the model describes an economy that is likely rife with inefficiencies arising from congestion and thick market externalities. Such a model does not lend itself well to the analysis of optimal, i.e. welfare maximizing, policies for the simple reason that it is not possible to disentangle the extent to which a particular tax schedule is "optimal" because it happens to correct inefficiencies particular to the Danish 1999-2003 labor market under consideration here, or because it optimally trades efficiency off with equity.[47]

In this last section of the paper, we do, however, take a step towards analyzing a notion of optimal taxation. Specifically, inspired by Werning (2007), Blundell and Shephard (2012) and Hosseini and Shourideh (2017), we use the estimated structural model to construct a Pareto optimal income tax reform. That is, we consider whether—taking the status quo in terms of the actual tax system and any inefficiencies that may be present in the baseline economy as given—it is possible to reform the income tax system in such a way that the government tax revenue increases while all workers are made (weakly) better off in terms of steady state expected utility.

To state the problem we solve formally, let $\overline{W}(a)$ be the expected steady state utility of a

[47] Bagger, Moen, and Vejlin (2017) provides a normative income taxation analysis in the context of a frictional labor market with two-sided heterogeneity.

type-a worker under the current tax system. We then solve

$$\max_{T(\cdot)} \int_0^1 \left\langle n^0(a)[T(b) - b] + [1 - n^0(a)] \int_b^{\overline{w}(a)} T(w) dG(w|a) \right\rangle dH(a) \tag{40}$$

subject to

$$n^0(a)\psi(c(a), s(b|a)) + [1 - n^0(a)] \int_b^{\overline{w}(a)} \psi(c(a), s(w|a)) dG(w|a) \geq \overline{W}(a) \tag{41}$$

for all $a \in [0, 1]$. At present, when searching for a Pareto optimal tax regime $T(\cdot)$ that solves (40) subject to (41), we retain the functional form of the actual tax system, but allow for different parameter values.[48]

The resulting Pareto improving tax system is superimposed on the actual 1999-2003 Danish income tax system in Figure 8. The reform increases government revenue by 0.68 percent while leaving no workers worse off. In fact, workers' steady state utility, on average, increases slightly by 0.07 percent. From Figure 8 we see that the Pareto optimal income tax regime reduces the marginal tax rate at low wage levels, and increases it slightly at all higher wage levels. At the same time, however, the average tax rate is increased at low wages, and reduced at higher wage levels.

Increasing the average tax rate at the bottom of the wage distribution, including in nonemployment, yields a mechanical revenue gain, while reducing the marginal tax rate at lower wage levels, predominantly relevant for low ability workers with high nonemployment rates, induces search at the bottom of the wage ladder, including among low ability nonemployed workers. This effect likely adds further government revenue. In order to leave workers at least as well of as under the baseline system, the average tax rate at higher wage levels must be reduced.

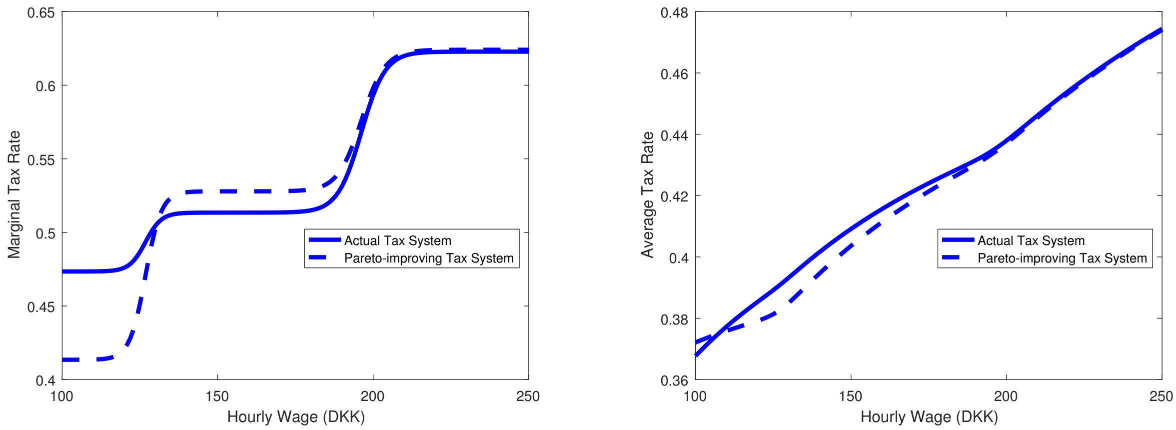
It is interesting to observe that the Pareto optimal tax schedule captures some qualitative features of the actual tax reforms implemented in Denmark during 1990-2005, and analyzed in section 6. In particular, tax rates at the very bottom of the wage distribution are increased, while they are lowered elsewhere. The Pareto optimal tax reform, however, combines these shifts in the tax function with a steeper marginal tax function, reducing marginal taxes for low wages and increasing it slightly elsewhere except at the very top. The actual reforms tended to uniformly reduce the marginal tax rates.

a-specific tax function $T(w; a)$. Finally, we discuss the possibility of conditioning the tax schedule on individual ability. In the model, ability a is observable to agents including the tax authority. Although the actual Danish tax system $T(w)$ does not condition on ability a, the social planner naturally has incentives to use an a-specific tax function $T(w; a)$ to improve welfare. This gives more freedom to the social planner, and we discuss potential results on an a-specific tax function $T(w; a)$.

We start this discussion by looking at the difference in several equilibrium variables between the actual tax system and the Pareto-improving tax system given in Figure 8. Tables 11 and 12

[48] An obvious alternative would be to look for a flexible, say, piece-wise linear tax system as in Blundell and Shephard (2012).

Figure 8: A Pareto optimal tax reform



calculate percentage changes of the following variables between these two tax regimes: aggregate unemployed search (S_0), aggregate employed search (S_1), aggregate hiring intensity (V), market tightness (θ), unemployment rate (n^0), hazard rate from non-employment to employment ($N2J$), job-to-job transition rate ($J2J$), and firms' wage policy (w^F). The baseline is the actual tax system (99-03). Table 11 is under partial equilibrium, while Table 12 is under full equilibrium. For example, in partial equilibrium and the Pareto-improving tax system, aggregate search intensity by unemployed S_0 is 0.95% higher than the actual tax system. From these two tables, one can find that behavior of low ability workers changes by a large amount due to the Pareto-optimal reform. This implies that the social planner finds room for welfare improvement among low ability workers and targets them. On the other hand, the opposite story holds for high ability workers; that is, they do not respond to the Pareto-improving tax reform. This is perhaps because the social planner does not care about them, who are already matched with good jobs under positive sorting.

The above observation implies that if the social planner can use an a-specific tax function $T(w;a)$, then the tax function for high ability workers would be similar to the actual tax function or Pareto-improving unconditional tax function especially around top brackets. They are matched with good jobs and thus there is no benefits from lowering top tax rates. In order to keep their utility, higher top tax rates are not feasible. Hence, top tax rates would be similar to status-quo tax rates. For low ability workers, one could see larger responses in the a-specific tax function $T(w;a)$, which will improve the equilibrium allocation and then bring higher utility and revenues. We are currently working on a-specific tax functions $T(w;a)$ in this Pareto-optimal tax reform exercise.

9 Conclusion

Income taxation reduces the return to job search effort, thus distorting workers' job search effort, and therefore the equilibrium allocation of heterogenous workers across heterogenous firms.

Table 11: Percentage change - Partial Equilibrium - Optimal tax system compared to 9903

α-type	S_0	S_1	V	θ	n^0	$N2J$	$J2J$	w^F
All	0.95	-0.17	0.00	0.00	-1.37	1.96	-0.12	0.00
0.1	3.17	1.38	0.00	0.00	-0.10	0.48	0.11	0.00
0.25	2.53	5.04	0.00	0.00	-5.44	7.47	0.91	0.00
0.3	1.35	0.91	0.00	0.00	-4.09	5.17	0.07	0.00
0.5	0.12	-0.96	0.00	0.00	-0.79	0.87	-0.42	0.00
0.6	0.04	-0.72	0.00	0.00	-0.29	0.31	-0.32	0.00
0.75	0.01	-0.25	0.00	0.00	-0.12	0.13	-0.12	0.00
0.9	0.00	-0.15	0.00	0.00	-0.06	0.06	-0.07	0.00

Table 12: Percentage change - Full equilibrium - Optimal tax system compared to 9903

α-type	S_0	S_1	V	θ	n^0	$N2J$	$J2J$	w^F
All	2.20	-1.14	0.75	-0.29	-3.11	4.62	-0.89	0.10
0.1	4.36	1.74	0.55	-0.01	-0.13	0.41	-0.12	0.01
0.25	3.54	4.09	-0.20	-2.61	-6.67	9.41	0.35	0.47
0.3	1.31	-6.79	0.61	0.83	-4.04	4.78	-2.75	0.04
0.5	-0.05	-3.31	0.87	1.85	-0.75	0.37	-1.74	0.01
0.6	-0.06	-1.44	0.75	1.23	-0.31	-0.05	-0.94	0.04
0.75	0.00	-0.24	0.57	0.66	-0.10	-0.18	-0.41	0.03
0.9	0.02	-0.19	0.54	0.61	-0.02	-0.25	-0.37	0.01

Using a rich equilibrium frictional labor market model of search on-the-job, we have documented economically significant effects of income taxation on the allocation of labor. Specifically, we found that 1994, 1999, and 2004 Danish income tax reforms improved the equilibrium allocation of labor, reducing the steady state unemployment rate by about 6 percent, and increasing steady state labor income by about 2 percent, with workers of low ability gaining the most in terms of lower unemployment rates and higher labor incomes.

We also computed a mobility-based long-run elasticity of taxable labor income to be 0.12, with the labor income of lower ability workers being more responsive to income taxation than that of high ability workers. A decomposition of the elasticity reveals that the larger response of low ability workers is mostly along the extensive margin, whereas the smaller adjustments towards the top of the ability distribution are along the intensive margin and the wage policy margin. Finally, we have identified a Pareto improving labor income taxation reform that increases government revenue by 0.7 percent while leaving all workers weakly better off. Throughout the paper we have emphasized that the distinctions between partial and full equilibrium, income and substitution effects, as well as worker heterogeneity are quantitatively important.

This paper has dealt exclusively with positive aspects of labor income taxation in a framework with labor market frictions and two-sided heterogeneity where the allocation of labor is endogenous with respect to the tax system. The model set up here does not lend itself easily to normative analysis, but in a related paper Bagger, Moen, and Vejlin (2017) studies optimal taxation in a similar environment, within a model better suited for normative analysis.

References

ABOWD, J. M., R. H. CREECY, AND F. KRAMARZ (2002): "Computing Person and Firm Effects Using Linked Longitudinal Employer-Employee Data," Longitudinal Employer-Household Dynamics Technical Papers 2002-06, Center for Economic Studies, U.S. Census Bureau.

ABOWD, J. M., F. KRAMARZ, AND D. N. MARGOLIS (1999): "High Wage Workers and High Wage Firms," *Econometrica*, 67(2), 251–333.

ACEMOGLU, D. (2009): *Introduction to Modern Economic Growth*. MIT Press.

ADAM, S., T. BESLEY, R. BLUNDELL, S. BOND, R. CHOTE, M. G. P. JOHNSON, G. MYLES, AND J. M. POTERBA (2011): *Dimension of Tax Design*. Oxford University Press.

ALTONJI, J. G., AND L. M. SEGAL (1996): "Small-Sample Bias in GMM Estimation of Covariance Structures," *Journal of Business and Economic Statistics*, 14, 353–366.

BAGGER, J., AND R. LENTZ (2015): "An empirical model of wage dispersion with sorting," .

BAGGER, J., E. R. MOEN, AND R. M. VEJLIN (2017): "Optimal Taxation with On-the-Job Search," .

BLUNDELL, R., AND T. MACURDY (1999): *Labour Supply: A Review of Alternative Approaches*vol. 3 of *Handbook of Labor Economics*, chap. 27. Elsevier Inc.

BLUNDELL, R., AND A. SHEPHARD (2012): "Employment, Hours of Work and the Optimal Taxation of Low Income Families," *Review of Economic Studies*, 79(2), 481–510.

BONTEMPS, C., J.-M. ROBIN, AND G. J. VAN DEN BERG (1999): "An Empirical Equilibrium Job Search Model with Search on the Job and Heterogeneous Workers and Firms," *International Economic Review*, 40(4), 1039–1074.

——— (2000): "Equilibrium search with continuous productivity dispersion: Theory and nonparametric estimation," *International Economic Review*, 41(2), 305–358.

BREDA, T., L. HAYWOOD, AND H. WANG (2017): "Labor Market Responses to Payroll Tax Reductions," .

BURDETT, K. (1978): "A Theory of Employee Job Search and Quit Rates," *American Economic Review*, 68(1), 212–220.

BURDETT, K., AND M. COLES (2003): "Equilibrium Wage Tenure Contracts," *Econometrica*, 71(5), 1377–1404.

——— (2010): "Wage/tenure contracts with heterogeneous firms," *Journal of Economic Theory*, 145(4), 1408–1435.

Burdett, K., and K. L. Judd (1983): "Equilibrium Price Dispersion," *Econometrica*, 51(4).

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.

Chetty, R., J. Friedman, T. Olsen, and L. Pistaferri (2011): "Adjustment Costs, Firm Responses, and Micro vs. Macro Labor Supply Elasticities: Evidence from Danish Tax Records," *Quarterly Journal of Economics*, 126, 749–804.

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.

Coles, M. (2001): "Equilibrium Wage Dispersion, Firm Size, and Growth," *Review of Economic Dynamics*, 4, 159–187.

Diamond, P. A. (1971): "A Theory of Price Adjustment," *Journal of Economic Theory*, 3, 156–168.

Feldstein, M. (1995): "The Effect of Marginal Tax Rates on Taxable Income: A Panel Study of the 1986 Tax Reform Act," *Journal of Political Economy*, 103(3), 551–572.

——— (1999): "Tax Avoidance and the Deadweight Loss of the Income Tax," *Review of Economics and Statistics*, 81(4).

Gentry, W. M., and R. G. Hubbard (2004): "The Effects of Progressive Income Taxation on Job Turnover," *Journal of Public Economics*, 88, 2301–2322.

Gourieroux, C., A. Monfort, and E. Renault (1993): "Indirect Inference," *Journal of Applied Econometrics*, 8, S85–S118.

Hall, R. E., and A. B. Krueger (2012): "Evidence on the Incidence of Wage Posting, Wage Bargaining, and On-the-Job Search," *American Economic Journal: Macroeconomics*, 4(4), 56–67.

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

Heckman, J. (1993): "What Has Been Learned About Labor Supply in the Past Twenty Years?," *American Economic Review*, 83(2), 116–121.

Hosseini, R., and A. Shourideh (2017): "Retirement Financing: An Optimal Reform Approach," .

Jakobsen, K., K. Jakobsen, H. Kleven, and G. Zucman (2018): "Wealth Taxation and Wealth Accumulation: Theory and Evidence from Denmark," Discussion paper, NBER Working Paper 24371.

Kennan, J. (2006): "A Note on Discrete Approximation of Continous Distributions," .

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.

le Maire, D., and B. Schjerring (2013): "Tax Bunching, Income Shifting and Self-Employment," *Journal of Public Economics*, 107, 1–18.

Lentz, R. (2010): "Sorting by Search Intensity," *Journal of Economic Theory*, 145(4), 1436–1452.

——— (2015): "Optimal Employment Contracts with Hidden Search," .

Lopes de Melo, R. (2018): "Firm Wage Differentials and Labor Market Sorting: Reconciling Theory and Evidence," *Journal of Political Economy*, 126(1), 313–346.

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

Melitz, M. J. (2003): "The impact of trade on intraindustry reallocations and aggregate industry productivity," *Econometrica*, 71(6), 1695–1725.

Merz, M. (1995): "Search in the labor market and the real business cycle," *Journal of Monetary Economics*, 36(2), 269–300.

Michau, J.-B. (2017): "On the Provision of Insurance Against Search-Induced Wage Fluctuations," .

Mirrlees, J., S. Adam, T. Besley, R. Blundell, S. Bond, R. Chote, M. G. P. Johnson, G. Myles, and J. M. Poterba (2011): *Tax by Design*. Oxford University Press.

Mortensen, D. T. (2005): *Wage Dispersion: Why Are Similar Workers Paid Differently?* MIT Press.

Piketty, T., and E. Saez (2013): *Optimal Labor Income Taxation*vol. 5 of *Handbook of Public Economics*, chap. 7, pp. 391–474. Elsevier Inc.

Postel-Vinay, F., and J.-M. Robin (2002): "Equilibrium wage dispersion with worker and employer heterogeneity," *Econometrica*, 70(6), 2295–2350.

——— (2004): "To match or not to match? To Match or Not to Match? Optimal wage policy with endogenous worker search intensity," *Review of Economic Dynamics*, 7, 297–330.

Saez, E., J. Slemrod, and S. H. Giertz (2012): "The Elasticity of Taxable Income with Respect to Marginal Tax Rates: A Critical Review," *Journal of Economics Literature*, 50(1), 3–50.

Shephard, A. (2017): "Equilibrium Search and Tax Credit Reform," *International Economic Review*, 58(4), 1047–1088.

Shimer, R. (2010): *Labor Markets and Business Cycles.* Princeton University Press.

Stevens, M. (2004): "Wage-Tenure Contracts in a Frictional Labour Market: Firms Strategies for Recruitment and Retention," *Review of Economic Studies*, 71(2), 535–551.

Werning, I. (2007): "Pareto Efficient Income Taxation," .

Appendix

A Proofs and derivations

A.1 Law of motions for $n_t^0(a)$ and $n_t^1(w|a)$

Recall from the main text that $n_t^0(a)$ is the time-t measure of unemployed workers in the household, while $n_t^1(w|a)$ is the measure of workers employed with wage w. Define

$$N_t(a) \equiv n_t^0(a) + \int_{\phi(a)}^{\overline{w}(a)} n_t^1(w|a)dw \tag{A1}$$

as the time-t measure of the household population.

Now, consider a small interval of time of length dt from t to $t+dt$. The change in the type-a household-specific unemployment rate from t to $t+dt$ is given by the difference in gross worker flows in and out of unemployment. Indeed,

$$n_{t+dt}^0(a) - n_t^0(a) = \delta(a)dt[N_t(a) - n_t^0(a)] - \lambda dt\,[\mu + s_t^0(a)]\overline{F}(\phi(a)|a)n_t^0(a). \tag{A2}$$

There are $N_t(a)-n_t^0(a)$ employed workers at time t, and within a small interval of time of length dt, a share $\delta(a)dt$ of them have their job destroyed, generating a gross flow $\delta(a)dt[N_t(a)-n_t^0(a)]$ into unemployment. There are $n_t^0(a)$ unemployed workers in the type-a household at time t, and between t and $t+dt$, a share $\lambda dt\,[\mu+s_t^0(a)]\overline{F}(\phi(a)|a)$ of them obtain an acceptable job offer, generating a gross flow $\lambda dt\,[\mu + s_t^0(a)]\overline{F}(\phi(a)|a)n_t^0(a)$ out of unemployment.

In equilibrium, $\overline{F}(\phi(a)|a) = 1$ as no firm would optimally set the offered type-a worker wage below the type-a household's reservation wage. Imposing $\overline{F}(\phi(a)|a) = 1$ and dividing through by $dt > 0$ and taking the limit as $dt \to 0$ yields the law of motion (6) stated in the main text.

The law of motion for $n_t^1(w|a)$ derives from a similar logic. Consider the change in the measure of workers employed at a job paying w or less during a small interval of time of length d from t to $t+dt$. That is, let $N_t^1(w|a) = \int_{\phi(a)}^{w} n_t^1(w|a)dw$, and consider

$$\begin{aligned} &N_{t+dt}^1(w|a) - N_t^1(w|a) = \\ &\quad \lambda dt[\mu + s_t^0(a)][F(w|a) - F(\phi(a)|a)]n_t^0(a) + \lambda dt\mu[F(w|a) - F(\phi(a)|a)][N_t(a) - n_t^0(a) - N_t^1(w|a)] \\ &\quad - [\delta(a) + \lambda\mu F(\phi(a)|a)]dtN_t^1(w|a) - \lambda dt\mu\overline{F}(w|a)N_t^1(w|a) - \overline{F}(w|a)\int_{\phi(a)}^{w} \lambda dts_t^1(x|a)n_t^1(x|a)dx. \end{aligned} \tag{A3}$$

The first term stems from the gross inflow of unemployed workers who find an acceptable job paying w or less. The second term stems from the inflow of workers currently employed at a wage greater than w (there are $N_t(a)-n_t^0(a)-N_t^1(w|a)$ such workers in a type-a household), but who are reallocated to an acceptable job paying w or less. The sum of the first two terms constitutes total inflow into the group of employed type-a household members earning a wage w or below. The third term reflects the gross worker outflow made up of employed workers earning a wage w or below who move into unemployment. They may do so due to a job destruction shock, or a reallocation shock involving an unacceptable job offer, an event that occurs with probability $\lambda\mu F(\phi(a)|a)dt$ in a short interval of length dt. The fourth term comprises gross worker outflow from workers earning a wage w or less who are reallocated to jobs paying in excess of w. Finally, the fifth term reflects gross outflow of workers who, through on-the-job search, located a job paying in excess of w. The sum of the last three terms constitutes total outflow from the group of employed type-a household members earning a wage w or below.

Rearranging and collecting terms in (A3), imposing the equilibrium restriction $F(\phi(a)|a) = 0$, dividing through by $dt > 0$ and taking the limit as $dt \to 0$ yields

$$\begin{aligned} \dot{N}_t^1(w|a) = {}& [\lambda\mu + \lambda s_t^0(a)]F(w|a)n_t^0(a) + \lambda\mu F(w|a)[N_t(a) - n_t^0(a) - N_t^1(w|a)] \\ & - \delta(a)N_t^1(w|a) - \lambda\mu\overline{F}(w|a)N_t^1(w|a) - \overline{F}(w|a)\int_{\phi(a)}^{w} \lambda s_t^1(x|a)n_t^1(x|a)dx. \end{aligned} \tag{A4}$$

Finally, taking the derivative with respect to w and a bit of algebra yields the law of motion (7) stated in the main text.

A.2 Steady state labor allocation

The size of a household is $N_t(a) = n_t^0(a) + \int_{\psi(a)}^{\overline{w}} n_t^1(w|a)$. Hence, $n_t^0(a)/N_t(a)$ is the type-a household time-t unemployment rate. The PDF of employed workers across wage rates is $g_t(w|a) = n_t^1(w|a)/[1 - n_t^0(a)]$, with associated CDF $G_t(w|a) = \int_{\phi(a)}^{\overline{w}} g_t(w|a)dw = N_t^1(w|a)/[1 - n_t^0(a)]$. The law of motions for $n_t^0(a)$ and $n_t^1(w|a)$ are discussed above. $G_t(w|a)$ is the steady state employment-weighted wage distribution, which reflects the allocation of labor across firms. The law of motion for $G_t(w|a)$ is straightforward to obtain from (A4). Indeed,

$$\dot{G}_t(w|a) = [\lambda\mu + \lambda s_t^0(a)]F(w|a)\frac{n_t^0(a)}{1 - n_t^0(a)} + \lambda\mu F(w|a)\overline{G}_t(w|a) \\ - \delta(a)G_t(w|a) - \lambda\mu\overline{F}(w|a)G_t(w|a) - \overline{F}(w|a)\int_{\phi(a)}^{w} \lambda s_t^1(x|a)g_t(x|a)dx. \quad \text{(A5)}$$

In steady state $\dot{n}_t^0(a) = 0$, and $s_t^0(a) = s^0(a)$, and $\dot{n}_t^1(w|a) = 0$ and $s_t^1(w|a) = s^1(w|a)$ for all $w \in [\psi(a), \overline{w}(a)]$, in which case (A2) implies that the worker-type conditional steady state unemployment rate, $n^0(a)$, is

$$n^0(a) = \frac{\delta(a)}{\delta(a) + \lambda\mu + \lambda s^0(a)}. \quad \text{(A6)}$$

Furthermore, in steady state, $\dot{G}_t(w|a) = 0$, in which case (A5) simplifies as

$$\int_{\phi(a)}^{w} \lambda s^1(x|a)g(x|a)dx = [\delta(a) + \lambda\mu]\frac{F(w|a) - G(w|a)}{\overline{F}(w|a)}, \quad \text{(A7)}$$

where we have used that, according to (A6), $n^0(a)/[1 - n^0(a)] = \delta(a)/[\lambda\mu + \lambda s^0(a)]$. Next, take the derivative of (A7) with respect to w to obtain

$$\lambda s^1(w|a)g(w|a) = [\delta(a) + \lambda\mu]\frac{f(w|a)\overline{G}(w|a) - g(w|a)\overline{F}(w|a)}{\overline{F}(w|a)^2}, \quad \text{(A8)}$$

which, upon multiplying through by $\overline{F}(w|a)/\overline{G}(w|a)$, yields

$$\lambda s^1(w|a)\overline{F}(w|a)\frac{g(w|a)}{\overline{G}(w|a)} = [\delta(a) + \lambda\mu]\frac{f(w|a)}{\overline{F}(w|a)} - [\delta(a) + \lambda\mu]\frac{g(w|a)}{\overline{G}(w|a)}. \quad \text{(A9)}$$

Rearrange to get an expression for the hazard function of $G(w|a)$,

$$\frac{g(w|a)}{\overline{G}(w|a)} = \left[\frac{\delta(a) + \lambda\mu}{\delta(a) + \lambda\mu + \lambda s^1(w|a)\overline{F}(w|a)}\right]\frac{f(w|a)}{\overline{F}(w|a)}. \quad \text{(A10)}$$

It follows that the steady state labor allocation, $G(w|a)$, is given by

$$G(w|a) = 1 - \exp\left\langle -\int_{\phi(a)}^{w}\left[\frac{\delta(a) + \lambda\mu}{\delta(a) + \lambda\mu + \lambda s^1(x|a)\overline{F}(x|a)}\right]\frac{f(x|a)}{\overline{F}(x|a)}dx\right\rangle. \quad \text{(A11)}$$

As mentioned in the main text, this expression is useful for the numerical solution of the model because it allow us to compute the allocation of labor from any search effort function $s^1(w|a)$ and wage offer distribution $F(x|a)$ without having to (numerically) solve a system of balanced flow equations.

The special case of wage-independent search effort, $s^1(w|a) = s^1$. While useful, the derivation of, and the resulting expression for, the employment-weighted wage distribution$G(w|a)$ is somewhat unconventional. Consider the special case where on-the-job search effort is independent of the wage. This exogenous search effort specification is common in the job search literature. Under the restriction that $s^1(w|a) = s^1(a)$, the integrated hazard function of the employment-weighted wage distribution is given by

$$\begin{aligned} \int_{\phi(a)}^{w}\frac{g(x|a)}{\overline{G}(x|a)}dx &= \int_{\phi(a)}^{w}\left[\frac{\delta(a) + \lambda\mu}{\delta(a) + \lambda\mu + \lambda s^1(a)\overline{F}(x|a)}\right]\frac{f(x|a)}{\overline{F}(x|a)}dx \\ &= \int_{\phi(a)}^{w}[1 + \kappa^1(a)\overline{F}(x|a)]^{-1}\frac{f(x|a)}{\overline{F}(x|a)}dx \\ &= \int_{\overline{F}(w|a)}^{1}[1 + \kappa^1(a)z]^{-1}\frac{1}{z}dz \\ &= \int_{\overline{F}(w|a)}^{1}\left[\frac{1}{z} - \frac{\kappa^1(a)}{1 + \kappa^1(a)z}\right]^{-1}dz. \end{aligned} \quad \text{(A12)}$$

where $\kappa^1(a) \equiv \frac{\lambda s^1(a)}{\delta(a)+\lambda\mu}$. Integrating (A12) yields

$$\begin{aligned}
\int_{\phi(a)}^{w} \frac{g(x|a)}{\overline{G}(x|a)}dx &= \Big|_{\overline{F}(w|a)}^{1} \ln\left(\frac{z}{1+\kappa^1(a)z}\right) \\
&= \ln\left(\frac{1}{1+\kappa^1(a)}\right) - \ln\left(\frac{\overline{F}(w|a)}{1+\kappa^1(a)\overline{F}(w|a)}\right) \\
&= \ln\left(\frac{1+\kappa^1(a)\overline{F}(w|a)}{[1+\kappa^1(a)]\overline{F}(w|a)}\right), \qquad \text{(A13)}
\end{aligned}$$

Hence, using (A11) on (A13), we obtain

$$G(w|a) = 1 - \exp\left\langle -\ln\left(\frac{1+\kappa^1(a)\overline{F}(w|a)}{[1+\kappa^1(a)]\overline{F}(w|a)}\right)\right\rangle = 1 - \frac{[1+\kappa^1(a)]\overline{F}(w|a)}{1+\kappa^1(a)\overline{F}(w|a)} = \frac{F(w|a)}{1+\kappa^1(a)\overline{F}(w|a)}. \qquad \text{(A14)}$$

We note that (A14) is the "standard" expression for the employment-weighted distribution of wages that appears in e.g. Burdett and Mortensen (1998) and Bontemps, Robin, and van den Berg (2000).

A.3 Applying the Maximum Principle

The current value Hamiltonian associated with the household problem is given as

$$\begin{aligned}
\mathscr{H}_t(a) &= \psi(c_t(a), s_t^0(a))n_t^0(a) + \int_{\phi(a)}^{\overline{w}(a)} \psi(c_t(a), s_t^1(w|a))n_t^1(w|a)dw \\
&\quad + e^{\rho t}\xi_t^0(a)\Big\langle \delta(a)N_t(a) - [\delta(a)+\lambda\mu+\lambda s_t^0(a)]n_t^0(a)\Big\rangle \\
&\quad + \int_{\phi(a)}^{\overline{w}(a)} e^{\rho t}\xi_t^1(x|a)\Big\langle\Big[\lambda\mu N_t(a) + \lambda s_t^0(a)n_t^0(a) + \int_{\phi(a)}^{x} \lambda s_t^1(z|a)n_t^1(z|a)dz\Big]f(x|a) \\
&\qquad - \Big[\delta(a)+\lambda\mu+\lambda s_t^1(x|a)\overline{F}(x|a)\Big]n_t^1(x|a)\Big\rangle dx, \qquad \text{(A15)}
\end{aligned}$$

with $c_t(a)$ given by (4), restated here for later reference,

$$c_t(a) = \frac{1}{N_t(a)}\left([b-T(b)]n_t^0(a) + \int_{\phi(a)}^{\overline{w}(a)} [w-T(w)]n_t^1(w|a)dw\right) + \overline{b} \qquad \text{(A16)}$$

and $N_t(a)$ given by (A1).

By the Maximum Principle (Acemoglu, 2009, Theorem 7.13, p. 254), the optimal controls satisfies the following necessary conditions. First, with respect to employed job search effort $s_t^1(w|a)$ at any wage level $w \in [\phi(a), \overline{w}(a)]$, optimality requires

$$\begin{aligned}
\frac{\partial \mathscr{H}_t(a)}{\partial s_t^1(w|a)} &= 0 \quad \text{for all } t\in[0,\infty) \text{ and for all } w\in[\phi(a),\overline{w}(a)], \qquad \text{(A17)} \\
\frac{\partial \mathscr{H}_t(a)}{\partial n_t^1(w|a)} &= -e^{\rho t}\dot{\xi}_t^1(w|a) \quad \text{for all } t\in[0,\infty) \text{ and for all } w\in[\phi(a),\overline{w}(a)], \qquad \text{(A18)}
\end{aligned}$$

where $\mathscr{H}_t(a)$ is given by (A15) and $\dot{\xi}_t^1(w|a) \equiv \partial\xi_t^1(w|a)/\partial t$. Of course, there is set of analogous necessary optimality conditions pertaining to unemployed search effort $s_t^0(a)$. Indeed,

$$\begin{aligned}
\frac{\partial \mathscr{H}_t(a)}{\partial s_t^0(a)} &= 0 \quad \text{for all } t\in[0,\infty), \qquad \text{(A19)} \\
\frac{\partial \mathscr{H}_t(a)}{\partial n_t^0(a)} &= -e^{\rho t}\dot{\xi}_t^0(a) \quad \text{for all } t\in[0,\infty). \qquad \text{(A20)}
\end{aligned}$$

We next solve the system of equation made up (A19), (A20), (A17), and (A18), with $\mathscr{H}_t(a)$ given by (A15).

Optimal steady state job search effort, $s_t^1(w|a)$ and $s_t^0(a)$. Consider first employed job search effort, $s_t^1(w|a)$. Using (A15), equating the derivative of $\mathscr{H}_t(a)$ with respect to $s_t^1(w|a)$, for some $w\in[\phi(a),\overline{w}(a)]$, to zero, and rearranging, yields

$$e^{-\rho t}\zeta'(s_t^1(w|a)) = \lambda\int_{w}^{\overline{w}(a)} [\xi_t^1(x|a) - \xi_t^1(w|a)]f(x|a)dx, \qquad \text{(A21)}$$

for any $w \in [\phi(a), \overline{w}(a)]$, and where $\zeta'(\cdot)$ represents the marginal disutility of job search effort, see (2). The necessary optimality conditions (A21) has a straightforward interpretation that is spelled out in the main text. For future reference, notice that integrating (A21) by parts implies

$$e^{-\rho t}\zeta'(s_t^1(w|a)) = \lambda \int_w^{\overline{w}(a)} \frac{\partial}{\partial x}\xi_t^1(x|a)\overline{F}(x|a)dx. \tag{A22}$$

Next, consider unemployed job search effort $s_t^0(a)$. Derivations analogous to those that lead to (A21) implies that

$$e^{-\rho t}\zeta'(s_t^0(a)) = \lambda \int_{\phi(a)}^{\overline{w}(a)} [\xi_t^1(x|a) - \xi_t^0(a)]f(x|a)dx, \tag{A23}$$

and integration by parts yields

$$e^{-\rho t}\zeta'(s_t^0(a)) = \lambda \int_{\phi(a)}^{\overline{w}(a)} \frac{\partial}{\partial x}\xi_t^1(x|a)\overline{F}(x|a)dx. \tag{A24}$$

The steady state costate $\xi_t^1(x|a)$. Consider first the costate variable $\xi_t^1(x|a)$, the multiplier on the law of motion for the control $n_t^1(w|a)$, the time-t measure of workers employed at wage w in the type-a household. Utilizing (A18) and equating the derivative of $\mathscr{H}_t(a)$ with respect to the control $n_t^1(w|a)$ to $-e^{\rho t}\dot{\xi}_t^1(w|a)$ yields

$$\begin{aligned} -\dot{\xi}_t^1(w|a) = e^{-\rho t}\Big[u(c_t(a)) - \zeta(s_t^1(w|a)) + u'(c_t(a))[w - T(w) - c_t(a)]\Big] \\ + \delta(a)[\xi_t^0(a) - \xi_t^1(w|a)] + \lambda\mu \int_{\phi(a)}^{\overline{w}(a)} [\xi_t^1(x|a) - \xi_t^1(w|a)]f(x|a)dx \\ + \lambda s_t^1(w|a) \int_w^{\overline{w}(a)} [\xi_t^1(x|a) - \xi_t^1(w|a)]f(x|a)dx, \end{aligned} \tag{A25}$$

for any $w \in [\phi(a), \overline{w}(a)]$. In deriving (A25) we have used that $\partial N_t(a)/\partial n_t^1(w|a) = 1$, and that, by (A16),

$$\frac{\partial c_t(a)}{\partial n_t^1(w|a)} = \frac{w - T(w) - c_t(a)}{N_t(a)}. \tag{A26}$$

The first term on the right-hand side of (A25) is the change in the household's flow-payoff if a worker earning a wage w is added to the household at time t. Such an additional worker increases the time-t utility flow in the household by $u(c_t(a)) - \zeta(s_t^1(w|a))$ and, in addition, contributes net-of-tax revenue in the amount of $w - T(w) - c_t(a)$ to be distributed within the household. The household values this additional revenue at the marginal utility of consumption, $u'(c_t(a))$. The two last terms on the right-hand side of (A25) represent the change in the value of the stock of workers earning a wage w, denominated in present values, stemming from the additional worker. This change is given by the expected capital gains from job destruction, reallocation and job search at the dictated search effort, $\lambda s_t^1(w|a)$ that the additional worker is subject to. At the optimal search policy, the change in the current flow-payoff plus the change to the value of the household's stock of workers employed at wage w, $n_t^1(w|a)$, balance the depreciation of the stock, i.e. $-\dot{\xi}_t^1(w|a)$.

Substituting (A21) into (A25) and rearranging allow us to write

$$-\dot{\xi}_t^1(w|a) = e^{-\rho t}\Phi_t^1(w|a) - [\delta(a) + \lambda\mu]\xi_t^1(w|a), \tag{A27}$$

where $\Phi_t^1(w|a)$ is defined as

$$\begin{aligned} \Phi_t^1(w|a) \equiv u(c_t(a)) - \zeta(s_t^1(w|a)) + u'(c_t(a))[w - T(w) - c_t(a)] \\ + \delta(a)e^{\rho t}\xi_t^0(a) + \lambda\mu \int_{\phi(a)}^{\overline{w}(a)} e^{\rho t}\xi_t^1(x|a)f(x|a)dx + \zeta'(s_t^1(w|a))s_t^1(w|a). \end{aligned} \tag{A28}$$

Solving the non-homogenous first-order linear ordinary differential equation (ODE) (A27) for $\xi_t^1(w|a)$ yields

$$\xi_t^1(w|a) = e^{[\delta(a)+\lambda\mu]t}\xi_0^1(w|a) - e^{[\delta(a)+\lambda\mu]t}\int_0^t e^{-[\rho+\delta(a)+\lambda\mu]x}\Phi_x^1(w|a)dx. \tag{A29}$$

Consider now an economy in steady state, i.e. consider the situation where $\dot{n}_t^1(w|a) = 0$ for all $w \in [\phi(a), \overline{w}(a)]$, as well as $\dot{n}_t^0(a) = 0$, such that $n_t^1(w|a) = n_{t'}^1(w|a) = n^1(w|a)$ for all $w \in [\phi(a), \overline{w}(a)]$, $n_t^0(a) =$

$n^0_{t'}(a) = n^0(a)$, and where $s^1_t(w|a) = s^1_{t'}(w|a) = s^1(w|a)$ and $s^0_t(a) = s^0_{t'}(a) = s^0(a)$ for any two dates t and t'. The steady state version (A28) implies that $\Phi^1_t(w|a) = \Phi^1_{t'}(w|a) = \Phi^1(w|a)$ for any two dates t and t'. Moreover, in steady state, the *current value* costate variables are equalized, i.e. $\xi^1_0(w|a) = e^{\rho t}\xi^1_t(w|a)$ for all $w \in [\phi(a), \overline{w}(a)]$. Put differently, the *present value* costate variables differ only across time due to discounting.[49] The steady state version of (A29) therefore implies that

$$\xi^0_t(w|a) = \frac{e^{[\delta(a)+\lambda\mu)]t}}{1 - e^{[\rho+\delta(a)+\lambda\mu)]t}}\left[1 - e^{-[\rho+\delta(a)+\lambda\mu]t}\right]\frac{\Phi^1(w|a)}{\rho + \delta(a) + \lambda\mu}. \tag{A30}$$

The transversality condition for the costate $\xi^1_t(w|a)$ holds. Indeed, since

$$\lim_{t\to\infty}\frac{e^{[\delta(a)+\lambda\mu]t}}{1 - e^{[\rho+\delta(a)+\lambda\mu]t}} = -\frac{\delta(a) + \lambda\mu}{\rho + \delta(a) + \lambda\mu}\lim_{t\to\infty} e^{-\rho t} = 0, \tag{A31}$$

we have

$$\lim_{t\to\infty}\xi^1_t(w|a) = 0. \tag{A32}$$

We note that the steady state identity $\xi^1_0(w|a) = e^{\rho t}\xi^1_t(w|a)$ and (A30) implies that the household's steady state shadow valuation of an additional worker employed with wage rate w is

$$[\rho + \delta(a) + \lambda\mu]\xi^1_0(w|a) = \Phi^1(w|a), \tag{A33}$$

which, upon substituting the steady state version of (A28) into (A33), implies that

$$\begin{aligned}\rho\xi^1_0(w|a) &= u(c(a)) - \zeta(s^1(w|a)) + u'(c(a))[w - T(w) - c(a)]\\ &+ \delta(a)[\xi^0_0(a) - \xi^1_0(w|a)] + \lambda\mu\int_{\phi(a)}^{\overline{w}(a)}[\xi^1_0(x|a) - \xi^1_0(w|a)]f(x|a)dx + \zeta'(s^1(w|a))s^1(w|a),\end{aligned} \tag{A34}$$

where $c(a)$ is steady state household consumption, see (A16). Substituting (A22) back into (A34), using again the steady state identity $\xi^1_0(w|a) = e^{\rho t}\xi^1_t(w|a)$, and integrating the remaining integral by parts yield

$$\begin{aligned}\rho\xi^1_0(w|a) &= u(c(a)) - \zeta(s^1(w|a)) + u'(c(a))[w - T(w) - c(a)]\\ &+ \delta(a)[\xi^0_0(a) - \xi^1_0(w|a)] + \lambda\mu\int_{\phi(a)}^{\overline{w}(a)}\frac{\partial}{\partial x}\xi^1_0(x|a)\overline{F}(x|a)dx + \lambda s^1(w|a)\int_{w}^{\overline{w}(a)}\frac{\partial}{\partial x}\xi^1_0(x|a)\overline{F}(x|a)dx.\end{aligned} \tag{A35}$$

By the Envelope Theorem, the derivative of (A35) with respect to w implies that

$$\frac{\partial}{\partial w}\xi^1_0(w|a) = \frac{u'(c(a))[1 - T'(w)]}{\rho + \delta(a) + \lambda\mu + \lambda s^1(w|a)\overline{F}(w|a)}. \tag{A36}$$

From (A36), we note that $\frac{\partial}{\partial w}\xi^1_0(w|a) > 0$, and hence, that $\xi^1_0(w|a)$ is strictly increasing in w. This implies that the household dictates that employed workers accept an alternative job offer if and only if it pays a higher wage than the current job.

Substituting (A36) back into (A35), allow us to express the steady state costate $\xi^1_0(w|a)$ as

$$\begin{aligned}\rho\xi^1_0(w|a) &= u(c(a)) - \zeta(s^1(w|a)) + u'(c(a))[w - T(w) - c(a)]\\ &+ \delta(a)[\xi^0_0(a) - \xi^1_0(w|a)] + \lambda\mu\int_{\phi(a)}^{\overline{w}(a)}\frac{u'(c(a))[1 - T'(x)]\overline{F}(x|a)}{\rho + \delta(a) + \lambda\mu + \lambda s^1(x|a)\overline{F}(x|a)}dx\\ &+ \lambda s^1(w|a)\int_{w}^{\overline{w}(a)}\frac{u'(c(a))[1 - T'(x)]\overline{F}(x|a)}{\rho + \delta(a) + \lambda\mu + \lambda s^1(x|a)\overline{F}(x|a)}dx.\end{aligned} \tag{A37}$$

The interpretation of (A37) is described in the main text.

Finally, substituting (A36) into (A22) yields a useful characterization of the optimal steady state search effort among employed workers,

$$\zeta'(s^1(w|a)) = \lambda\int_{w}^{\overline{w}(a)}\frac{u'(c(a))[1 - T'(x)]\overline{F}(x|a)}{\rho + \delta(a) + \lambda\mu + \lambda s^1(x|a)\overline{F}(x|a)}dx. \tag{A38}$$

[49] Recall from the main text that $\xi^1_t(w|a)$ is the present value shadow value of adding an additional employed worker earning a wage w to the household at time t. In steady state, this differs from the present value shadow value of adding an additional employed worker at wage rate w to the household at time 0, $\xi^1_0(w|a)$, only by discounting. Hence, $\xi^1_0(w|a) = e^{\rho t}\xi^1_t(w|a)$.

The steady state costate $\xi^0_t(a)$. The characterization of the costate $\xi^0_t(a)$ is analogous to that of $\xi^1_t(w|a)$ detailed above, but is nonetheless included here for completeness. First, equate the derivative of $\mathscr{H}_t(a)$ with respect to $n^0_t(a)$ to $-e^{\rho t}\dot{\xi}^0_t(a)$, to obtain

$$-\dot{\xi}^0_t(a) = e^{-\rho t}\Big[u(c_t(a)) - \zeta(s^0_t(a)) + u'(c_t(a))[b - T(b) - c_t(a)]\Big] + [\lambda\mu + \lambda s^0_t(a)]\int_{\phi(a)}^{\overline{w}(a)} [\xi^1_t(x|a) - \xi^0_t(a)]f(x|a)dx. \quad \text{(A39)}$$

Equation (A39) has the same interpretation as (A25).

Substituting (A23) into (A39) and rearranging yields

$$-\dot{\xi}^0_t(a) = e^{-\rho t}\Phi^0_t(a) - \lambda\mu\xi^0_t(a), \quad \text{(A40)}$$

where $\Phi^0_t(a)$ is defined as

$$\Phi^0_t(a) \equiv u(c_t(a)) - \zeta(s^0_t(a)) + u'(c_t(a))[b - T(b) - c_t(a)] + \int_{\phi(a)}^{\overline{w}(a)} e^{\rho t}\xi^1_t(x|a)f(x|a)dx + \zeta'(s^0_t(a))s^0_t(a). \quad \text{(A41)}$$

The solution to the non-homogenous first-order linear ODE (A40) can be expressed as

$$\xi^0_t(a) = e^{\lambda\mu t}\xi^0_0(a) - e^{\lambda\mu t}\int_0^t e^{-[\rho+\lambda\mu]x}\Phi^0_x(a)dx. \quad \text{(A42)}$$

Imposing steady state, where $\Phi^0_t(a) = \Phi^0(a)$ and $\xi^0_0(a) = e^{\rho t}\xi^0_t(a)$, on (A42) implies

$$\xi^0_t(a) = \frac{e^{\lambda\mu t}}{1 - e^{[\rho+\lambda\mu]t}}\left[1 - e^{-[\rho+\lambda\mu]t}\right]\frac{\Phi^0(a)}{\rho + \lambda\mu}. \quad \text{(A43)}$$

It is straightforward to verify that the transversality condition $\lim_{t\to\infty}\xi^0_t(a) = 0$ holds, and that the household's steady state shadow valuation of an additional unemployed worker, $\xi^0_t(a)$ is

$$[\rho + \lambda\mu]\xi^0_0(a) = \Phi^0(a). \quad \text{(A44)}$$

Substituting the steady state version of (A41) into (A44) yields

$$\rho\xi^0_0(a) = u(c(a)) - \zeta(s^0(a)) + u'(c(a))[b - T(b) - c(a)] + \lambda\mu\int_{\phi(a)}^{\overline{w}(a)} [\xi^1_0(x|a) - \xi^0_0(a)]f(x|a)dx + \zeta'(s^0(a))s^0(a), \quad \text{(A45)}$$

where $c(a)$ is steady state household consumption, see (A16). Substitute (A24) into (A45) and integrate the integral by parts to obtain

$$\rho\xi^0_0(a) = u(c(a)) - \zeta(s^0(a)) + u'(c(a))[b - T(b) - c(a)] + \lambda[\mu + s^0(a)]\int_{\phi(a)}^{\overline{w}(a)} \frac{\partial}{\partial x}\xi^1_0(x|a)\overline{F}(x|a)dx. \quad \text{(A46)}$$

The derivative $\frac{\partial}{\partial x}\xi^1_0(x|a)$ is given by (A36), and hence, we may write

$$\rho\xi^0_0(a) = u(c(a)) - \zeta(s^0(a)) + u'(c(a))[b - T(b) - c(a)] + \lambda[\mu + s^0(a)]\int_{\phi(a)}^{\overline{w}(a)} \frac{u'(c(a))[1 - T'(x)]\overline{F}(x|a)}{\rho + \delta(a) + \lambda\mu + \lambda s^1(x|a)\overline{F}(x|a)}dx. \quad \text{(A47)}$$

Finally, substituting (A36) into (A24) yields,

$$\zeta'(s^0(a)) = \lambda\int_{\phi(a)}^{\overline{w}(a)} \frac{u'(c(a))[1 - T'(x)]\overline{F}(x|a)}{\rho + \delta(a) + \lambda\mu + \lambda s^1(x|a)\overline{F}(x|a)}dx. \quad \text{(A48)}$$

The steady state reservation wage, $\phi(a)$. The steady state reservation wage $\phi(a)$ solves

$$\xi^1_0(\phi(a)|a) = \xi^0_0(a). \quad \text{(A49)}$$

In other words, it is the wage rate $\phi(a)$ at which the household's shadow valuation of employment coincides with the shadow valuation of unemployment. From (A37), (A47), (A38) and (A48) it is evident that

$$\phi(a) = b. \quad \text{(A50)}$$

and that unemployed steady state search effort is identical to employed steady state search effort in the lowest paying firm with $w = b$,

$$s^0(a) = s^1(b|a). \quad \text{(A51)}$$

A.4 Proof of Lemma 1

Proof. Consider property (i) in lemma 1. From (13), $s(w|a) = [\zeta']^{-1}\left(\lambda \int_{w}^{\overline{w}(a)} \frac{u'(c(a))[1-T'(x)]\overline{F}(x|a)}{\rho+\delta(a)+\lambda\mu+\lambda s(x|a)\overline{F}(x|a)}dx\right) + \underline{s}$. Since $\zeta' : [0,\infty) \mapsto [0,\infty)$ we have $[\zeta']^{-1} : [0,\infty) \mapsto [0,\infty)$. Hence, $[\zeta']^{-1}(\cdot) \geq 0$, and therefore, $s(w|a) \geq \underline{s}$. Consider now property (ii) in lemma 1. The derivative of $[\zeta']^{-1}(\cdot)$ is $1/\zeta''([\zeta']^{-1}(\cdot))$. Since $\zeta''(\cdot) > 0$, $[\zeta']^{-1} : [0,\infty) \mapsto [0,\infty)$ is a strictly increasing function. It follows that $[\zeta']^{-1}(0) = 0$, and hence, $s(\overline{w}(a)|a) = 0$. Finally, consider property (iii) in lemma 1. $[\zeta']^{-1}(\cdot)$ is a strictly increasing function. Take $w' < w''$. Then, $\int_{w'}^{\overline{w}(a)} \frac{u'(c(a))[1-T'(x)]\overline{F}(x|a)}{\rho+\delta(a)+\lambda\mu+\lambda s(x|a)\overline{F}(x|a)}dx > \int_{w''}^{\overline{w}(a)} \frac{u'(c(a))[1-T'(x)]\overline{F}(x|a)}{\rho+\delta(a)+\lambda\mu+\lambda s(x|a)\overline{F}(x|a)}dx$. It follows that $s(w'|a) > s(w''|a)$.[50] □

A.5 Proof of Lemma 2

Proof. Take $w' < w''$ both in $[b, \overline{w}(a)]$. Lemma 1 established that $s(w'|a) > s(w''|a)$, from which it follows that $s(w'|a)\overline{F}(w'|a) > s(w''|a)\overline{F}(w''|a)$. Finally, $\int_b^{w'} s(x|a)dG(x|a) < \int_b^{w''} s(x|a)dG(x|a)$. Lemma 2 now follows from (21) and (22). □

A.6 Proof of Proposition 1

Proof. Let $p' < p''$ such that $y(a,p') < y(a,p'')$, and let $w' = w(a|p')$ and $w'' = w(a|p'')$. Then, $\tilde{\pi}(p''|w'',a) = [y(a,p'') - w'']\tilde{\ell}(a|w'') > [y(a,p'') - w']\tilde{\ell}(a|w') > [y(a,p') - w']\tilde{\ell}(a|w') = \tilde{\pi}(p'|w',a) > [y(a,p') - w'']\tilde{\ell}(a|w'')$. This implies that $[y(a,p'') - y(a,p')]\tilde{\ell}(a|w'') > [y(a,p'') - y(a,p')]\tilde{\ell}(a|w')$, which implies $\tilde{\ell}(a|w'') > \tilde{\ell}(a|w')$. Lemma 2 established that $\tilde{\ell}(a|w)$ is strictly increasing in w, which implies that $w' < w''$. □

A.7 Proof of Proposition 2

Proof. Fix $a \in [0,1]$. The first order condition (FOC) for a type-p firm's hiring intensity problem (23) is $\tilde{\pi}(a|p)h(a)/V(a) = d'(v)$. The left-hand side of the FOC is independent of v. Strict convexity of $d(\cdot)$ implies that the right-hand sider of the FOC is strictly increasing in v. Hence, the FOC has a unique solution $v = v(a|p)$. Strict convexity of $d(\cdot)$ also ensures the second order condition is satisfied. The proof of Proposition 1 established that $\tilde{\pi}(p|a)$ is strictly increasing in p. Strict convexity of $d(\cdot)$ implies that $v(a|p)$ is strictly increasing in p. □

A.8 Proof of Proposition 3

Proof. Taking the derivative of (23), applying the Envelope Theorem, yields

$$\frac{\partial \pi(p)}{\partial p} = (1-\tau)\int_0^1 \left[\frac{v(a|p)}{V(a)}\right]\frac{\partial \tilde{\pi}(p|a)}{\partial p}dH(a).$$

The proof of Proposition 1 established that $\tilde{\pi}(p|a)$ is strictly increasing in p for every $a \in [0,1]$. Proposition 2 established that $v(a|p)$ is strictly increasing in p for every $a \in [0,1]$. The product of strictly increasing strictly positive functions yields a strictly increasing function. Integrating a set of strictly increasing functions yields a strictly increasing function. Proposition 3 follows. □

B Details on the data

B.1 The merging procedure

Merging data sources by person ID is straightforward. However, a few complications arise when merging data sources by firm IDs because different public registers using different firm identifiers. Data clean-up, manipulation, merging and selection is carried out in a series of SAS-programs. Before we move on to describing the merging procedure, we provide a brief description of the person and firm IDs that enters the procedure.

[50] Alternatively, since $F(w|a)$ is differentiable, direct differentiation of (13) yields the desired result.

Person and firm IDs. Persons are identified by their CPR-number. All individuals residing legally in Denmark are registered in the CPR register (Centrale Person Register) with a unique CPR number. The CPR-number is the sole identifier of an individual in relation to the state. Person data is recorded under a anonymized CPR-number labeled `pnr`.

There are two main ways of identifying business entities. The Central Business Register (CVR), established in October 1999, contains primary data on all businesses with economic activity in Denmark, regardless of economic and organizational structure. CVR covers both public and private businesses.The SE-number identifies a business in Stamregistret for Erhvervsdrivende (the SE-register), a register established in 1985. The main function of SE-register is to identify businesses vis-a-vis the tax authorities when settling value added tax (VAT), income tax payments, or tax payments for self-employed persons, and pension contributions. CVR- and SE-numbers are recorded under the variable names `cvrnr` and `senr`, respectively. Although the relationship between CVR-numbers and SE-numbers is somewhat complicated by fact that a business entity may split up into multiple legal business entities, an SE-numbers associated with one and only one CVR-number at a given point in time. We are able to map CVR-numbers and SE-numbers using a correspondence table provided by Statistics Denmark.

As it turns out, employing firms in the labor market spell data are identified by the firm ID `spell_firm_id`, a hybrid firm identifier specific to the labor market spell data, which corresponds very closely to the CVR-number. As we describe below, we are able to map `spell_firm_id` to `cvrnr` using IDA-N data.

The labor market spell data. The unit of observation in the labor market spell data is a person-spell-year, where a spell is a job with a particular employer, and the employing firm is identified by `spell_firm_id` as described above. That is, a labor market spell that stretches across three calendar years (i.e. stretches across three January 1st) is represented by three observations in the labor market spell data.

Merging IDA-P onto the labor market spell data. We first merge the employment spell data and IDA-P. IDA-P contains annual background information from public registers on all individuals aged 15-74 residing legally in Denmark on the 31st of December. The unit of observation in IDA-P is a person-year. We merge on person identifier (variable `pnr`) and year (variable `aar`). 98.64% of the employment spell observations are matched to an IDA-P observation. We retain only person-years that are found in both data sources.

Merging IDA-S onto the labor market spell data. IDA-S contains background information from public records on all physical workplaces in Denmark, excluding the so-called fictitious workplaces briefly described in the main text. The unit of observation in the raw IDA-S panel is a workplace-year. Our analysis is conducted at the firm-level, and a firm may consist of several workplaces. Because IDA-S contains both workplace and firm identifiers, we can aggregate IDA-S to firm-years, letting a firm inherit industry affiliation and public sector status from its largest workplace.[51]

We start by mapping the firm identifier of employment spell data (variable `spell_firm_id`) into the firm identifier of the IDA database (cvrnr). The mapping between `spell_firm_id` and `cvrnr` is established using IDA-N, yet another IDA dataset. IDA-N contains data on all primary employment relationships ongoing on November 28th in a given year. The unit of observation in IDA-N is a person-workplace-year, but it also contains information on both the IDA-S firm identifier `cvrnr` and the spell data firm identifier (`spell_firm_id`), allowing us to construct a correspondence table with `spell_firm_id`, `cvrnr`, and `aar`. We use this table to map `spell_firm_id` into `cvrnr` in the labor market data,[52] and then we subsequently merge IDA-S onto the employment spells on firm identifier `cvrnr` and year `aar`. 96.10% of the employment spell observations are matched with an IDAS observation. We retain all employment spell data, whether matched to IDA-S or not.

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

[52]There is a few occurrences where the mapping between `spell_firm_id` and `cvrnr` is 1-to-many, in which case we arbitrarily pick `cvrnr` as max(`cvrnr`).

Merging VAT data onto the labor market spell data. As described in the main text, value added data comes from two data sources, MOMS and MOMM. MOMS covers 1995-1999 and contains annual data on firms' sales and purchases, whereas MOMM starts in January, 2000, and provides the same information, but recorded on a monthly frequency.[53]

The MOMS data is recorded under the `senr`-firm ID, whereas the labor market spell data (and MOMM) is recorded under the firm ID `cvrnr`. As a first step in the merging of VAT data to the labor market spell data, we map `senr` in MOMS to `cvrnr`. This done using a correspondence table provided by Statistics Denmark. Then, using MOMM data up until December, 2005, we aggregate the monthly data to annual frequencies. Stacking the MOMS and MOMM data now yields a value added panel covering the period 1995-2005, containing essentially all firms with economic activity in Denmark over the period. The unit of observation in this combined MOMS and MOMM data is a firm-year, with the firm ID given by `cvrnr`.

Merging the annual value added panel onto the labor market spell data, we match 90.37% of the employment panel firm-months to value added data.[54] We retain all employment spell data, whether matched to the VAT-data or not. This completes the merging procedure. As mentioned in the main text, the resulting 1990-2005 matched employer-employee panel contains 84,474,045 spell observations on 4,447,401 individuals, 428,448 firms, 57,568,393 job spells and 26,905,652 non-employment spells.

C Estimating $T'(\cdot)$ and $T(\cdot)$

Consider the 1999 - 2003 regime. Given that this regime has a progressive structure with three brackets (see Table 5), we impose the following flexible parametric function for marginal tax rates $T'(\cdot)$,

$$T'(w) = \frac{\alpha_1}{1+\exp(-\alpha_2(w-\alpha_3))} + \frac{\alpha_4}{1+\exp(-\alpha_5(w-\alpha_6))} + \alpha_7. \tag{C1}$$

The marginal tax function $T'(\cdot)$ is composed of two logistic functions "splined" together.

The data points we use the compute the parameters α_1, α_2, ..., α_7 consist of 100 hourly wage distribution percentiles and 100 median marginal tax rates for each of the 5 years. Let $(w_{tk}^{(50)}, [T']_{tk}^{(50)})$ for $t = 1999, 2000, ..., 2003$ and $k = 1, 2, ..., 100$ denote these data points. The parameters α_1, α_2, ..., α_7 are estimated by fitting $T'(w_i)$ to t_i, where $T'(\cdot)$ is given by (C1). That is,

$$\min_{\alpha_1,\alpha_2,...,\alpha_7} \sum_{t=1999}^{2003} \sum_{k=1}^{100} \left[\frac{\alpha_1}{1+\exp(-\alpha_2(w_{tk}^{(50)}-\alpha_3))} + \frac{a_4}{1+\exp(-\alpha_5(w_{tk}^{(50)}-\alpha_6))} + \alpha_7 - [T']_{tk}^{(50)} \right]^2. \tag{C2}$$

Let $\hat{\alpha}_1$, $\hat{\alpha}_2$, ..., $\hat{\alpha}_7$ be the solution to the minimization problem (C2).

Integrating (C1) produces the following expression for the average tax function $T(\cdot)$,

$$T(w) = \frac{\alpha_1}{\alpha_2} \log(e^{\alpha_2\alpha_3} + e^{\alpha_2 w}) + \frac{\alpha_4}{\alpha_5} \log(e^{\alpha_5\alpha_6} + e^{\alpha_5 w}) + \alpha_7 w + \alpha_0, \tag{C3}$$

where α_0 is the constant of integration. Parameter estimates for α_1, α_2, ..., α_7 are available, and we estimate α_0 in the same way we estimated α_1, α_2, ..., α_7. That is, we calculate median tax liability $T_{tk}^{(50)}$, measured in the same unit as w_{tk}, for each of the 100 wage percentiles and for each of the years 1999 - 2003, and fit (C3) with $\alpha_1 = \hat{\alpha}_1$, $\alpha_2 = \hat{\alpha}_2$, ..., $\alpha_7 = \hat{\alpha}_7$ to $(w_{tk}, T_{tk}^{(50)})$ with $t = 1999, 2000, ..., 2003$ and $k = 1, 2, ..., 100$:

$$\min_{\alpha_0} \sum_{t=1999}^{2003} \sum_{k=1}^{100} \left[\frac{\hat{\alpha}_1}{\hat{\alpha}_2} \log(e^{\hat{\alpha}_2\hat{\alpha}_3} + e^{\hat{\alpha}_2 w_{tk}}) + \frac{\hat{\alpha}_4}{\hat{\alpha}_5} \log(e^{\hat{\alpha}_5\hat{\alpha}_6} + e^{\hat{\alpha}_5 w_{tk}}) + \hat{\alpha}_7 w_{tk} + \alpha_0 - T_{tk}^{(50)} \right]^2. \tag{C4}$$

[53] The VAT declaration frequency depends on the firm's annual turnover. Firms register VAT information at the tax authorities monthly if their turnover exceeds DKK 15 mill., quarterly if their turnover is between DKK 1 mill. and DKK 15 mill., and every six months if turnover is below DKK 1 mill. The reporting frequency for a given firm is also provided in the MOMM data.

[54] The share of observations in the value added panel matched to the employment panel is much lower due to many firms in the value added data being small firms with no employees.

The marginal and average tax functions pertaining to the other tax regimes under consideration in this paper, i.e. the 1990-1993, the 1994-1998, and the 2004-2005 tax regimes, are estimated by analogous procedures, sometimes adapted to account for specific features of a tax regime. For example, for the 1990-1993 tax regime, and for the years 1994 and 1995, we build the marginal tax functions from three, rather than two, logistic functions.

D Numerical implementation

D.1 Solving the model

D.1.1 Discretization of $H(a)$ and $\Gamma_0(p)$

Our algorithm is based on discrete values of worker ability a and firm productivity p, whereas the theoretical model takes the distributions $H(a)$ and $\Gamma_0(p)$ to be continuously differentiable. We apply Proposition B of Kennan (2006) to discretize the support of $H(\cdot)$. Let N_a and N_p be the numbers of discrete points in the support of worker ability a and firm productivity p, respectively.[55] For $i = 1, ..., N_a$, $a_i = H^{-1}([2i-1]/2N_a)$, where $H(\cdot)$ is the (continuous) population distribution of a.[56] We then calculate a probability mass function (PMF) $\hat{h}(a_i)$ and a corresponding cumulative distribution function (CDF) $\hat{H}(a_i)$ as $\hat{h}(a_i) = 1/N_a$ and $\hat{H}(a_i) = i/N_a$, introducing the "ˆ"-notation to distinguish the continuous population distributions and their discrete approximations. The same procedure applies for p, where we compute $\hat{\gamma}_0(p_i) = 1/N_p$ and $\hat{\Gamma}_0(p_i) = i/N_p$.

D.1.2 Finding $\lambda(a)$, $v(a|p)$, and $w(a|p)$

This is the main part of numerical implementation. Given $\hat{H}(a)$ and $\hat{\Gamma}_0(p)$, the structure of the algorithm is as follows.

0. Guess $\lambda_0(a)$, $s(w|a)$, $v_0(a|p)$, and $w_0(a|p)$
1. Calculate $s_1(w|a)$, $w_1(a|p)$, $v_1(a|p)$, and $\lambda_1(a)$
2. If $s_1(w|a)$, $w_1(a|p)$, $v_1(a|p)$, and $\lambda_1(a)$ differ from $s_0(w|a)$, $w_0(a|p)$, $v_0(a|p)$, and $\lambda_0(a)$, take $s_0(w|a) = s_1(w|a)$, $w_0(a|p) = w_1(a|p)$, $v_0(a|p) = v_1(a|p)$, and $\lambda_0(a) = \lambda_1(a)$ and go to 1. If not, take $s(w|a) = s_1(w|a)$, $w(a|p) = w_1(a|p)$, $v(a|p) = v_1(a|p)$, and $\lambda(a) = \lambda_1(a)$ and terminate the algorithm.

We explain calculations of step 1 in detail below.

Calculation of $F(w(a|p)|a)$. The first step is to calculate $F(w(a|p)|a)$. For the current guesses of $\lambda(a)$, $v(a|p)$, and $w(a|p)$, from (28), the discrete approximation of $F(w(a_i|p_j)|a_i)$ is given by

$$\hat{F}(w(a_i|p_j)|a_i) = \frac{\sum_{z=p_0}^{p_j} v(a_i|z)\hat{\gamma}_0(z)}{\sum_{z=p_0}^{1} v(a_i|z)\hat{\gamma}_0(z)}.$$

The associated PMF $\hat{f}(w(a_i|p_j)|a_i)$ is calculated as $\hat{F}(w(a_i|p_j)|a_i) - \hat{F}(w(a_i|p_{j-1})|a_i)$.

Calculation of $s(w|a)$. For the current guesses of $\lambda(a)$, $v(a|p)$, and $w(a|p)$, we employ the following algorithm to compute $s(w|a)$.

0. Guess $s(w_j|a_i) = s_0$.
1. Calculate $n^0(a_i)$ from (14) and $\hat{G}(w_j|a_i)$ from (16).
2. Calculate s_1 as the RHS of (13), replacing $F(w|a)$ with $\hat{F}(w_j|a_i)$, and approximating the integral as detailed further below.
3. If $s_1 \neq s_0$, take $s_0 = s_1$ and go to 1. If $s_1 = s_0$, take $s(w_j|a_i) = s_1$ and terminate the algorithm.

[55] In the empirical implementation we take $N_a = 100$ and $N_p = 1,000$.

[56] Recall that $H(\cdot) = \mathscr{U}[0,1]$ in our empirical implementation.

Step 2 involves solving for $s(w_j|a_i)$ as the fixed point of (13). We approximate the integral in (13) as follows:

$$\begin{aligned}
\int_{w_{ij}}^{\overline{w}(a_i)} \frac{u'(c(a_i))[1-T'(x)]\overline{F}(x|a_i)}{\rho+\delta(a_i)+\lambda\mu+\lambda s(x|a_i)\overline{F}(x|a_i)}dx &= \int_{w_{ij}}^{w_{iN_p}} \phi_i(x)dx \\
&= \int_{w_{ij}}^{w_{iN_p}} \phi_i(x)\frac{1}{f(x|a_i)}dF(x|a_i) \\
&= \sum_{k=j}^{N_p} \phi_i(w_{ik})\frac{1}{f(w_{ik}|a_i)}\hat{f}(w_{ik}|a_i) \\
&= \sum_{k=j}^{N_p-1} \phi_i(w_{ik})\frac{1}{f(w_{ik}|a_i)}\hat{f}(w_{ik}|a_i) \\
&\approx \sum_{k=j}^{N_p-1} \phi_i(w_{ik})\frac{\hat{f}(w_{ik}|a_i)}{\hat{f}(w_{i,k+1}|a_i)}(w_{i,k+1}-w_{ik}),
\end{aligned}$$

where $\phi_i(x) \doteq \frac{u'(c(a_i))[1-T'(x)]\overline{F}(x|a_i)}{\rho+\delta(a_i)+\lambda\mu+\lambda s(x|a_i)\overline{F}(x|a_i)}$. For the second equality above, we use the definition of a PDF, i.e., $dF(x|a_i) = f(x|a_i)dx$, while the third equality comes from the definition of expectation in discrete w. The fourth equality is from the fact that $\phi_i(w_{iN_p}) = 0$ for all i. We finally use numerical differentiation to approximate $f(w_{ik}|a_i) \approx \frac{F(w_{i,k+1}|a_i)-F(w_{ik}|a_i)}{w_{i,k+1}-w_{ik}} = \frac{\hat{f}(w_{i,k+1}|a_i)}{w_{i,k+1}-w_{ik}}$ for $k \leq N_p - 1$. Note that $s(w_{iN_p}|a_i) = s(\overline{w}(a_i)|a_i) = 0$ for all i.[57]

D.1.3 Calculate $w(a|p)$

From (24), $w(a_i|p_j)$ is calculated as

$$w(a_i|p_j) = y(a_i,p_j) - \frac{\tilde{\pi}(a_i|p_j)}{\tilde{\ell}(a_i|w_{i\cdot j})}.$$

$\tilde{\ell}(a_i|w_{i\cdot j})$ is calculated from (22). We approximate $\tilde{\pi}(a_i|p_j)$ by the second order Taylor series sequentially:

$$\tilde{\pi}(a_i|p_{j+1}) \approx \tilde{\pi}(a_i|p_j) + \frac{\partial\tilde{\pi}(a_i|p_j)}{\partial p}(p_{j+1}-p_j) + \frac{1}{2}\cdot\frac{\partial^2\tilde{\pi}(a_i|p_j)}{\partial p^2}(p_{j+1}-p_j)^2. \tag{D1}$$

By the Envelope Theorem,

$$\frac{\partial\tilde{\pi}(a_i|p_j)}{\partial p} = \frac{\partial y(a_i,p_j)}{\partial p}\tilde{\ell}(a_i|w_{i\cdot j}).$$

By taking the second derivative and applying numerical differentiation,

$$\begin{aligned}
\frac{\partial^2\tilde{\pi}(a_i|p_j)}{\partial p^2} &= \frac{\partial^2 y(a_i,p_j)}{\partial p^2}\tilde{\ell}(a_i|w_{ij}) + \frac{\partial y(a_i,p_j)}{\partial p}\cdot\frac{\partial\tilde{\ell}(a_i|w_{ij})}{\partial w}\cdot\frac{\partial w_{ij}}{\partial p} \\
&\approx \frac{\partial^2 y(a_i,p_j)}{\partial p^2}\tilde{\ell}(a_i|w_{ij}) + \frac{\partial y(a_i,p_j)}{\partial p}\cdot\frac{\tilde{\ell}(a_i|w_{i,j+1})-\tilde{\ell}(a_i|w_{ij})}{w_{i,j+1}-w_{ij}}\cdot\frac{w_{i,j+1}-w_{ij}}{p_{j+1}-p_j}.
\end{aligned}$$

Thus, by substituting these partial derivatives, we can calculate (D1) as

$$\tilde{\pi}(a_i|p_{j+1}) \approx \tilde{\pi}(a_i|p_j) + \frac{\partial y(a_i,p_j)}{\partial p}\cdot\frac{\tilde{\ell}(a_i|w_{ij})+\tilde{\ell}(a_i|w_{i,j+1})}{2}(p_{j+1}-p_j) + \frac{\partial^2 y(a_i,p_j)}{\partial p^2}\cdot\frac{\tilde{\ell}(a_i|w_{ij})}{2}(p_{j+1}-p_j)^2.$$

The initial condition is given by $\tilde{\pi}(a_i|p_0) = 0$ for all a_i from $\pi(p_0) = 0$.

D.1.4 Calculate $v(a|p)$

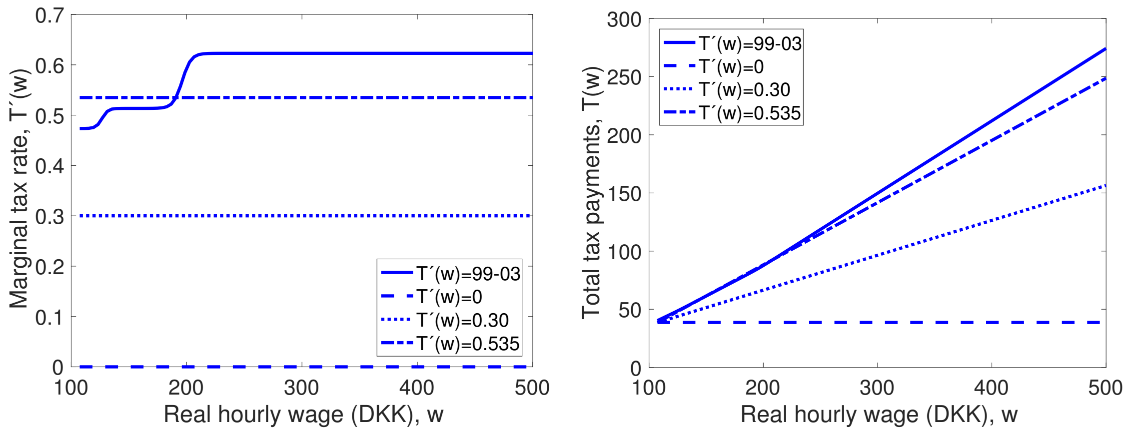
M_0 is given by (27), which gives $V(a) = M_0\int_{p_0}^{1} v(p)d\Gamma_0(p)$. We calculate $v(a|p)$ from (23).

D.1.5 Calculate $\lambda(a)$

$S(a)$ is given by (29), which gives $\lambda(a) = (V(a)/S(a))^{1-\eta}$.

[57] We compared $s(w_{ij}|a_i)$ calculated from the method explained here with the one calculated from value function iteration, and confirmed that two methods gave the same results with negligible approximation errors. Due to computational time, our method is preferred to value function iteration.

Figure E.1: Counterfactual linear tax systems



Panel A: $T'(\cdot)$ Panel B: $T(\cdot)$

E Allocative distortions and tax progressivity

Our evaluations of the 1990, 1994, 1999, and 2004 Danish tax reforms as well as the analysis of the elasticity of taxable income suggest that income taxations induce quantitative relevant distortions in the equilibrium allocation of labor. The Danish income tax schedule is characterized by a high degree of progressivity, see Figures 7, and in this section we consider a number of counter-factual reforms to quantify the impact of income tax progressivity on the allocation of labor.

Specifically, we consider three counterfactual linear tax regimes, with $T'(w) = 0.54$, $T'(w) = 0.30$, and $T'(w) = 0$. The $T'(w) = 0.54$ is the mechanically revenue neutral linear tax regime, that is, in the absence of any behavioral changes, replacing the actual 1999-2003 income tax system with a the linear tax function $T'(w) = 0.54$ would leave the government budget unchanged. The $T'(w) = 0.30$ tax regime is intended as a brute approximation of the US income tax system, and the extreme counterfactual $T'(w) = 0$ tax system is included primarily for illustration. When we impose each of the counterfactual tax systems, we always impose that $b - T(b)$ stays constant. The counterfactual tax regimes are rendered graphically in Figure E.1 along with the actual 1999-2003 income tax system.

Table E.1 present the percentage changes in nonemployment, labor income and the government budget for each of the three counterfactual tax regimes. As usual, we report both partial and full equilibrium responses, and also report the compensated responses where we keep the marginal utility of consumption constant. We stress again, the we do not balance the government budget in these cases, an issue that may be particularly egregious in the context of large counterfactual reforms where the direct effect on the government budget may be rather large, and one might suspect that income effects arising from lump sum transfers balancing the budget may be quantitatively important.[58]

The mechanically revenue neutral linear tax $T'(w) = 0.54$ increases nonemployment, and reduced expected labor income in both full and partial equilibrium almost everywhere in the ability distribution, except at the very top. In the aggregate, the $T'(w) = 0.54$ regime therefore induces an increase in the nonemployment rate, a decrease in aggregate labor income, as well as a small reduction in the government budget stemming from changes in the behavior of workers and firm.

Moving on to the $T'(w) = 0.30$ regime, we see relatively large reduction in the nonemployment rate, especially in partial equilibrium, and fairly large increases in labor income, across the ability distribution. This implies that the aggregate nonemployment rate is reduced by almost 10 percent in full equilibrium, while aggregate labor

[58] As mentioned above, we are currently working on amending our computation algorithms to impose balanced government budgets, but are at present unable to present empirical results for this case.

Table E.1: The impact of income tax progressivity on nonemployment, labor income and the government budget

		$\frac{\Delta n^0}{n^0} \times 100\%$				$\frac{\Delta LI}{LI} \times 100\%$				$\frac{\Delta B}{B} \times 100\%$			
		Total		Subst		Total		Subst		Total		Subst	
Comparison tax regime	Ability	PE	FE	PE	FE	PE	FE	PE	FE	PE	FE	PE	FE
$T'(w) = 0.54$	All	2.9	5.7	2.3	5.5	−0.4	−0.6	−0.3	−0.4	−1.0	−2.1	−0.8	−1.9
	0.25	8.4	14.4	7.8	15.2	−2.6	−5.6	−2.4	−5.9	−5.8	−14.5	−5.2	−15.5
	0.50	4.5	5.6	4.9	6.9	−0.5	−1.2	−0.5	−1.5	2.1	1.2	2.0	0.8
	0.75	0.9	2.4	−0.1	1.1	0.1	1.4	0.1	1.7	−1.2	0.1	0.0	1.3
$T'(w) = 0.30$	All	−10.7	−5.0	−18.3	−9.8	1.6	3.1	2.4	5.0	−24.2	−24.1	−22.8	−23.1
	0.25	−13.9	−10.2	−17.1	−9.7	4.2	4.5	5.2	4.4	4.7	2.2	8.0	1.6
	0.50	−12.8	−5.3	−20.4	−9.9	1.4	2.0	2.3	3.4	−15.4	−16.1	−14.0	−14.9
	0.75	−12.0	−3.3	−26.2	−14.6	1.0	3.2	1.9	5.9	−24.5	−24.2	−23.5	−22.7
$T'(w) = 0$	All	−19.6	−11.8	−32.3	−20.5	2.9	5.7	4.4	9.2	−56.8	−57.4	−55.2	−58.4
	0.25	−29.0	−24.6	−34.4	−25.3	8.8	11.0	10.5	11.4	2.3	−2.9	7.6	2.3
	0.50	−23.8	−12.1	−37.2	−21.6	2.7	4.1	4.2	7.2	−40.6	−43.0	−38.7	−42.2
	0.75	−20.1	−6.5	−41.6	−24.2	1.3	3.5	3.1	9.1	−57.0	−58.6	−56.0	−58.7

income increases by 5 percent. The impact on the government budget is very large, with the full equilibrium effect being a 24 percent drop in government revenue. We caution that if we were to make workers absorb this reduction in the budget through lump-sum taxation, it would likely induce rather large income effects in job search behavior. These could alter the conclusion drawn regarding the effect of going to a $T'(w) = 0.30$ regime on nonemployment and labor income. The same caveats apply to the interpretation of the last counterfactual exercise presented in Table E.1, where we impose the extreme tax regime $T'(w) = 0$, while keeping $b - T(b)$ constant. This reform has the same qualitative predictions as discussed above for the case of $T'(w) = 0.30$ regime, but quantitatively, the impact is of course large, and that includes a very large and negative impact on the government budget.