Tilburg University

The optimal taxation of unskilled labor with job search and social assistance

Boone, J.; Bovenberg, A.L.

Publication date:
2004

Link to publication

THE OPTIMAL TAXATION OF UNSKILLED LABOR
WITH JOB SEARCH AND SOCIAL ASSISTANCE

Jan Boone
Lans Bovenberg

The Optimal Taxation of Unskilled Labor with Job Search and Social Assistance
Jan Boone and Lans Bovenberg
NBER Working Paper No. 9785
June 2003
JEL No. H2, J2

ABSTRACT

In order to explore the optimal taxation of low-skilled labor, we extend the standard model of optimal non-linear income taxation in the presence of quasi-linear preferences in leisure by allowing for involuntary unemployment, job search and an exogenous welfare benefit. In trading off low-skilled employment against work effort of higher skilled workers, the government balances distortions on the search margin with those on work effort. Higher welfare benefits typically reduce taxes paid by low-skilled workers and raise marginal tax rates throughout the skill distribution.

Jan Boone
Department of Economics
Tilburg University
P.O. Box 90153
5000 LE Tilburg
The Netherlands
j.boone@uvt.nl

Lans Bovenberg
CentER
Tilburg University
P.O. Box 90153
5000 LE Tilburg
The Netherlands
a.l.bovenberg@uvt.nl

The optimal taxation of unskilled labor with job search and social assistance [*]

Jan Boone[†]and Lans Bovenberg[‡]

21-04-2003

Abstract

In order to explore the optimal taxation of low-skilled labor, we extend the standard model of optimal non-linear income taxation in the presence of quasi-linear preferences in leisure by allowing for involuntary unemployment, job search and an exogenous welfare benefit. In trading off low-skilled employment against work effort of higher skilled workers, the government balances distortions on the search margin with those on work effort. Higher welfare benefits typically reduce taxes paid by low-skilled workers and raise marginal tax rates throughout the skill distribution.

J.E.L. codes: H2, J2

Key words: labor-market search, social assistance, unemployment, low-skilled labor, non-linear income taxation, participation margin.

1 Introduction

Widening wage dispersion raises the question how public policy should protect the living standards of unskilled workers, as policy makers are increas-

[†]CentER, TILEC, Tilburg University and CEPR, London. email: j.boone@uvt.nl. Address: Dept. of Economics, Tilburg University, P.O. Box 90153, 5000 LE Tilburg, The Netherlands.

[‡]CentER, TILEC, Tilburg University and CEPR, London. email: a.l.bovenberg@uvt.nl. Address: CentER, Tilburg University, P.O. Box 90153, 5000 LE Tilburg, The Netherlands.

ingly concerned about the adverse incentive effects of generous income support. In response to these concerns, many countries have cut taxes on unskilled work in order to combat poverty while at the same time encouraging unskilled workers to look for work. Both the United States and several European countries have already introduced or are considering in-work tax benefits for unskilled work in the form of an Earned Income Tax Credit (EITC). These tax policies are part of active labor-market policies and 'welfare-to-work' programs, where governments fight poverty by raising employment of unskilled workers.

To investigate the optimal response of tax policy to income support provided through the welfare system and to declining relative wages of unskilled labor, we extend the standard model of optimal non-linear income taxation developed by Mirrlees (1971). In particular, we incorporate labor-market imperfections that induce governments to provide income support, namely search costs and involuntary unemployment.

In the presence of a search margin, the government has to account for not only the standard incentive compatibility constraint on work effort, but also a participation constraint on the willingness of low-skilled agents to look for work. Indeed, from an analytical point of view, our main contribution is to add a participation constraint to the optimal tax problem, including the decision regarding which types should optimally participate in job search. In doing so, we extend both the optimal tax literature, which typically abstracts from participation constraints, and the literature on optimal non-linear monopoly pricing, which generally assumes that the lowest participating type is exogenously given. Within a non-linear pricing framework, Rochet and Stole (2002) recently added an endogenous participation constraint by allowing agents to differ in both outside options and preferences for quality. Our analysis deviates from that of Rochet and Stole (2002) in two respects. First, in the non-linear pricing problem explored in Rochet and Stole (2002), the monopolist cares only about profits earned on the participating agents. In our optimal tax problem, in contrast, also agents who do not participate appear in the objective function because the government is interested in the utilities of both participating and nonparticipating types. The second difference with Rochet and Stole (2002) is that we allow agents to differ in only one dimension; agents feature different skill levels but exhibit the same search costs. Within the context of our labor-market application, this is a reasonable assumption, which is in fact employed by most of the labor-market literature on search (see, for instance, Mortensen and Pissarides (1999)). This assumption implies that the participation constraint is binding only at the bottom of the skill distribution. Rochet and Stole (2002), in contrast, derive a binding participation constraint for each

2

type.

The literature on optimal income taxation has modelled unemployment of unskilled agents as these agents reducing the hours they work in their jobs when they face low gross wages and rapidly rising marginal tax rates. Accordingly, low productivity workers are bunched in low- or zero-production jobs. By introducing a participation margin and positive search costs, we model another type of bunching at the bottom of the skill distribution: unskilled agents do not search for work and thus drop out of the labor force. Heckman (1993), for instance, stressed that 'a crucial theoretical distinction with important empirical pay off is that between labor supply choices at the extensive margin (...) and choices at the intensive margin'. Empirical work does in fact reveal that unskilled workers adjust their labor supply in response to tax and benefit programs on mainly the extensive margin (i.e. leaving the labor force altogether, for example through early retirement) rather than the intensive margin (i.e. reducing the hours they work in their jobs) (see, e.g. Eissa and Liebman (1996), Kimmel and Kniesners (1998), Blundell (2001), Meyer and Rosenbaum (2001) and Meyer (2002)). This explains the policy concern about welfare programs and high taxes on unskilled work discouraging low-income earners from looking for work. Indeed, our model is consistent with the stylized fact that low-skilled agents feature the highest long-term unemployment rates (see OECD (2001)).

Also Saez (2002) incorporates the two labor-supply margins of not only hours worked but also labor-force participation in an optimal income tax model.[1] Our approach differs from that of Saez in three important respects. First of all, whereas Saez assumes that all unemployed have voluntarily left the labor force, we account also for involuntary unemployment. Whereas allowing for unemployment risk does not complicate the analysis, it provides an implicit justification for substantial welfare benefits. Agents thus face two risks: being born with low ability and being involuntarily unemployed. More generally, we are more explicit than Saez (2002) about the labor-market imperfections affecting the costs and effectiveness of labor-market search, including the welfare implications of these imperfections.

Our analysis differs from Saez (2002) also in that the government takes the welfare benefit as exogenously given when optimizing the tax system. Hence, the government can employ only the non-linear income tax to optimize social welfare. Indeed, in practice, taxes and social assistance are often set

---

[1]Diamond (1980) and Choné and Laroque (2002) incorporate a participation margin in an optimal tax framework in which work effort is exogenous so that the intensive margin is absent. Our paper is similar to Saez (2002) in that we model both the intensive and extensive margins of labor supply, allowing us to explore the interaction between these two labor-supply margins.

by distinct agencies based on rather different interests and considerations. In some federal countries, for example, local governments determine social assistance benefits, while the central government is mainly responsible for the tax system. One can also interpret the minimum income floor set by social assistance benefits as being determined by considerations outside our model. Alternatively, one can view our analysis as exploring how the tax system can be employed to address the possibly sub-optimal aspects of social assistance.

A final difference is that Saez (2002) allows for more general preferences that are not necessarily quasi-linear in leisure. Whereas his results are thus more general than ours, our specific assumptions on preferences enable us to derive more analytical results on comparative statics with respect to public spending, labor-market imperfections (such as the costs and effectiveness of search) and institutional features of the welfare system. This sheds additional light on the determinants of the optimal tax schedule. Indeed, a substantial literature (see, e.g., Boadway, Cuff and Marchand (2000), Ebert (1992), Weymark (1986, 1987), and Lollivier and Rochet (1983)) has turned to quasi-linear preferences in leisure in order to obtain more intuition for the determinants of the optimal non-linear income tax, as these preferences allow for closed-form solutions of the standard optimal non-linear income tax problem. Our quasi-linear preferences also imply that a utilitarian government cares about the distribution of consumption rather than the distribution of work effort. Indeed, policy debates typically focus on raising consumption rather than reducing work effort of the poor.

We extend the literature on optimal non-linear income taxation with quasi-linear preferences in three ways; we allow for involuntary unemployment, a participation (or search) constraint, and an exogenous welfare benefit. These three extensions make this literature more relevant for addressing the timely policy question of how the optimal tax system should treat low-skilled employment in the face of income support. The participation margin also eliminates some of the unrealistic implications of a model with quasi-linear preferences for aggregate labor supply elasticities. Hence, the introduction of a participation margin is particularly important in the context of this particular model.

The rest of this paper is structured as follows. After section 2 introduces the model, section 3 sets out the optimal tax problem. In exploring the consequences of a binding participation margin, section 4 discusses, among other things, how labor-market imperfections and the features of the welfare system impact the optimal income tax. Section 5 concludes. The main results are proved in the appendix. The proofs of some auxiliary results can be found in Boone and Bovenberg (2002), which is the working paper version of this article.

2   The model

The economy is populated with agents featuring homogenous preferences but heterogeneous skills. A worker of ability (or skill or efficiency level) n working y hours (or providing y units of work effort) supplies ny efficiency units of homogeneous labor. With constant unitary labor productivity, these efficiency units are transformed in the same number of units of output. We select output as the numeraire. The before-tax wage per hour is thus given by exogenous skill n. Hence, overall gross output produced by a worker of skill n, $z(n)$, amounts to $z(n) = ny(n)$. Since workers collect only labor income, this gross output $z(n)$ corresponds to the gross (i.e. before-tax) labor income earned by a worker of that skill n. The density of agents of ability n is denoted by $f(n)$, and $F(n)$ represents the corresponding cumulative distribution function. The support of the distribution of abilities is given by $[n_0, n_1]$,[2] while $f(.)$ is differentiable and satisfies $1 - n_0 f(n_0) > 0$.

Workers share the following quasi-linear utility function over consumption x and hours worked (or work effort) y

$$u(x,y) = v(x) - y,$$

where $v(x)$ is increasing and strictly concave: $v'(x) > 0, v''(x) < 0$ for all $x \geq 0$. Furthermore, $v(0) = 0$, $\lim_{x \downarrow 0} v'(x) = \infty$, $\lim_{x \longrightarrow +\infty} v'(x) = 0$ and $\lim_{x \to +\infty} v(x) - xv'(x) = +\infty$. The concavity of $v(.)$ implies that agents are risk averse and thus want to obtain insurance against the risks of involuntary unemployment and a low earning capacity n. The specific cardinalization of the utility function affects the distributional preferences of a utilitarian government. In particular, the concavity of $v(.)$ implies that a utilitarian government aims to fight poverty. In other words, such a government wants to insure agents against the risk of a low consumption level.

As in Lollivier and Rochet (1983), Weymark (1987), Ebert (1992), and Boadway, Cuff and Marchand (2000), utility is linear in work effort y and separable in work effort and consumption x. This has four important consequences. First, consumption x is not affected by income effects. A higher average tax rate thus induces households to raise work effort y rather than to cut consumption x. Second, the single-crossing (or sorting) property is met, implying that the incentive compatibility constraints can be replaced by

---

[2]The finite upper bound of the skill distribution n_1 implies that the optimal marginal tax rate is zero at the top of the income distribution (see Sadka (1976) and Seade (1977)). Diamond (1998) considers various unbouded distributions, including the Pareto distribution for skills above the mode. We focus on the bottom rather than the top of the skill distribution and therefore do not extensively consider the sensitivity of our results with respect to the distribution of skills at the top of the skill distribution.

(much simpler) monotonicity conditions on $x(.)$ and $z(.)$ (see, for instance, Fudenberg and Tirole (1991)). Third, the specific quasi-linear utility function allows for a closed-form solution of the standard optimal income tax problem. Fourth, a utilitarian government cares only about aggregate work effort in the economy. Such a government thus aims at an equal distribution of consumption (i.e. the alleviation of poverty) rather than an equal distribution of work effort over the various agents. Indeed, Kanbur, Keen and Tuomala (1994) observe that policy debates focus on raising consumption rather than reducing work effort (or increasing leisure) of the poor. However, whereas Kanbur, Keen, and Tuomala (1994) adopt a non-welfarist social welfare function to do justice to this policy concern of fighting poverty, we continue in the welfarist tradition but assume a special, quasi-linear utility function.

In line with the optimal income tax literature, the government is assumed not to be able to observe skills n but to know the distribution function $f(n)$ and before-tax income of each individual $z(n)$. We depart from the standard optimal tax literature by incorporating non-verifiable job search: agents have to search for a job and the government cannot verify search intensities. In particular, we allow agents to adjust their labor supply on not only the intensive margin (i.e. by varying hours of work) but also the extensive margin (i.e. by deciding whether or not to look for a job). In particular, by searching with intensity $s \in [0,1]$, agents find a job with probability s. Search costs $\gamma(s)$ are given by

$$\gamma(s) = \begin{cases} \gamma s & \text{if} \quad s \in [0, \bar{s}] \\ +\infty & \text{otherwise,} \end{cases}$$

where $\gamma \geq 0$ is a parameter representing the magnitude of the search costs. $\bar{s} < 1$ captures the idea that agents can fail to find a job, even though they search at full capacity $\bar{s}$. By modelling the costs and effectiveness of search, the parameters γ and $(1 - \bar{s})$ represent labor-market imperfections that give rise to unemployment. Agents thus differ in both ability n and employment status and face two types of risks: being born with low ability n and being involuntarily unemployed.

If an agent does not succeed in finding a job, (s)he receives a welfare (or social assistance) benefit $b \geq 0$.[3] Since the government cannot observe

---

[3] An alternative interpretation of b is a categorial unemployment insurance benefit. Indeed, the benefit is paid only to those who have not found a job. In most countries, however, unemployment benefits depend on the previously earned wage incomes and are thus likely to increase with ability n. This is the main reason why we interpret b as a social assistance payment, i.e. the minimum income level provided by the government. Another interpretation of b is an early retirement or disability benefit that is paid if an agent does not have a job.

the abilities and search intensities of unemployed agents,[4] the welfare benefit does not depend on n and is exogenously given to the agent. An agent of ability n thus selects search intensity s to maximize expected utility

$$U(n) = \max_s \left\{ -\gamma(s) + su(n) + (1-s)\,v(b) \right\}$$

where $u(n)$ denotes the utility of having a job for a worker of type n. Substituting in the search cost function $\gamma(s)$ introduced above, one can easily verify that the optimal choice of s for type n amounts to

$$s(n) = \begin{cases} 0 & \text{if} \quad u(n) < \gamma + v(b) \\ \bar{s} & \text{if} \quad u(n) \geq \gamma + v(b). \end{cases} \tag{1}$$

The linear specification of the search cost function thus implies that a worker either does not search at all and is voluntarily unemployed or that he searches at the level $\bar{s}$ (and faces a probability of $(1-\bar{s})$ of involuntary unemployment). We refer to the constraint $u(n) \geq \gamma + v(b)$ as the participation or individual rationality constraint. The government has to respect this participation constraint because it cannot observe search. A special case of our model is a regular labor-supply model with fixed costs of work, e.g. child care costs. In particular, without involuntary unemployment (i.e. $\bar{s} = 1$), the parameter γ can be interpreted as these fixed costs of entering the labor market. Indeed, an agent enters the labor market by actively looking for a job only if the additional pay-off from work, $u(n) - v(b)$, exceeds the fixed entry cost γ. Positive entry costs introduce a difference between stopping to work on the intensive and extensive margins of labor supply: working zero hours in a job (i.e. responding on the intensive margin) is different from staying outside the labor force by not looking for a job (i.e. responding on the extensive margin). Unemployment risk (i.e. $\bar{s} < 1$) justifies substantial welfare benefits b. This generalization of the standard model comes at no analytical cost.

After a worker has found a job, (s)he has to determine her work effort. Instead of working with work effort $y(n)$ and consumption $x(n)$ as the instruments of the worker, we write the utility function in terms of gross income (or output) $z(n) \equiv ny(n)$ and net income (or consumption) $x(n)$. Utility of type n is then written as $u(n) \equiv v(x(n)) - z(n)/n$. The ex-post utility of a type n agent who finds a job is determined by type n's choice of gross income z:

$$u(n) = \max_z \left\{ v\left(z - \tilde{T}(z) \right) - \frac{z}{n} \right\}, \tag{2}$$

---

where $\tilde{T}(z)$ denotes the tax schedule as a function of gross income z. We can write $T(n) = \tilde{T}(z(n))$ since type n chooses gross income $z(n)$ in equilibrium. The envelope theorem yields the first-order incentive compatibility constraint

$$u'(n) = \frac{z(n)}{n^2}. \tag{3}$$

The government maximizes ex-ante expected utility (i.e. expected utility before the search effort has been made and therefore before labor market status has been revealed):

$$W \equiv \int_{n_0}^{n_1} \left[-\gamma s(n) + s(n)u(n) + (1 - s(n))v(b) \right] f(n)\phi(n)dn.$$

We normalize the rank-order weights $\phi(n)$ such that $\int_{n_0}^{n_1} f(n)\phi(n) = 1$, and assume $\phi'(n) \leq 0$.[5] The government is utilitarian if the rank-order weights are constant, i.e. $\phi(n) = 1$ for all n. This is the usual assumption adopted in the literature on optimal non-linear income taxation in the presence of preferences that are quasi-linear in leisure (see Lollivier and Rochet (1983), Weymark (1987), Ebert (1992), and Boadway, Cuff, and Marchand (2000)[6]). If the welfare weights are declining (i.e. $\phi'(n) < 0$), the government is concerned about the distribution of not only consumption but also leisure (or work effort).

The government has to respect the following budget constraint

$$\int_{n_0}^{n_1} f(n) s(n) [b + T(n)]dn = E + b, \tag{4}$$

where E represents exogenously given exhaustive government expenditure, and $T(n) \equiv z(n) - x(n)$ denotes the tax paid by type n. The government can employ only the non-linear income tax to optimize social welfare and takes public spending E and the welfare benefit as given. Note that the government's financial balance improves by $T(n) + b$ if a worker of type n finds a job. Alternatively, from the worker's point of view, $T(n) + b$ represents the implicit tax rate on entering the labor market by engaging in search. If $T(n) + b < 0$, entry is subsidized for type n.

---

[5]The rank-order weights depend on ability n rather than utility $u(n)$. This approach, which involves non-welfarists elements, allows us to derive a closed-form solution for the standard optimal tax problem. Atkinson (1995) defends this assumption by noting that empirical measures of inequality are based on the distribution of gross wages n rather than utilities.

[6]This paper considers also a maxi-min objective function where the government cares only about the least able persons (i.e. agents with skill n_0). This is the special case of our formulation in which $\phi(n) = 0$ for $n > n_0$.

3 The optimal tax problem

In optimizing social welfare, the government faces the incentive compatibility constraint (3), the participation constraint (1), and the government budget constraint (4). In addition, labor supply and therefore before-tax income should be non-negative

$$z(n) \geq 0. \tag{5}$$

Morover, the second-order condition for the agents' optimal choice of consumption and gross income implies that consumption and gross income are non-decreasing in type n (see, for instance, Fudenberg and Tirole (1991: theorem 7.2) or Boone and Bovenberg (2002: lemma 1))

$$z'(n) \geq 0. \tag{6}$$

Instead of $x(n)$, we employ $u(n)$ as a control variable in order to facilitate the inclusion of incentive compatibility (3) into our optimization problem.[7] To incorporate the monotonicity constraint (6), we introduce a non-negative variable $\omega(n) \equiv z'(n)$ determining how fast $z(n)$ rises with ability n. This yields the following optimization problem

$$\max_{\substack{s(.),u(.),z(.), \\ \omega(.) \geq 0}} \int_{n_0}^{n_1} \left\{ \begin{array}{c} [s(n)(u(n) - \gamma) + (1 - s(n))v(b)]\phi(n)f(n) - \lambda_u(n)\left[u'(n) - \frac{z(n)}{n^2}\right] \\ -\lambda_z(n)[z'(n) - \omega(n)] - \eta(n)(\gamma - u(n) + v(b)) \\ +\lambda_E f(n)[s(n)(T(n) + b) - b - E] - \delta(n)[0 - z(n)] \end{array} \right\} dn,$$

where $T(n) \equiv z(n) - x(n) = z(n) - v^{-1}\left(u(n) + \frac{z(n)}{n}\right)$. $\lambda_u(n)$ and $\lambda_z(n)$ represent the Lagrange multipliers of the first-order compatibility constraint and the monotonicity constraint, λ_E stands for the multiplier of the government budget constraint, $\eta(n)$ denotes the multiplier of the participation constraint, and $\delta(n)$ is the Lagrange multiplier of the non-negativity constraint on before-tax income (5).

To further simplify the optimization problem, we observe that, since $z(n) \geq 0$, incentive compatibility (3) implies that utilities do not decline with skill (i.e. $u'(n) \geq 0$). Accordingly, if the participation constraint $u(n) \geq \gamma + v(b)$ is met for skill $\bar{n}$, it is met also for all higher skills $n > \bar{n}$.

---

[7]This is the usual approach in optimal non-linear income tax problems introduced by Mirrlees (1971). If preferences are quasi linear in leisure, however, it is more convenient to employ $x(n)$ and $nu(n)$ rather than $z(n)$ and $u(n)$ as controls; see Boadway, Cuff, and Marchand (2000). We stick to the Mirrlees approach because we gain less by departing from it, due the introduction of a participation constraint on $u(n)$.

Defining n_w as the lowest skill looking for work, we thus have $s(n) = 0$ for $n < n_w$ and $s(n) = \bar{s}$ for $n \geq n_w$. Accordingly, our model is consistent with the stylized fact that the extensive margin of labor-supply responses is especially relevant for low-income earners. Empirical work does show that these workers respond to tax and benefit programs mainly on the extensive rather than the intensive margin of labor supply. This explains the policy concern that generous benefit programs and taxes on low earnings discourage low-income earners from looking for work.

The agents with skill $n < n_w$ can be viewed as being voluntarily unemployed. In our model in which all agents feature the same preferences, the lower skills $n < n_w$ are voluntarily unemployed, whereas the higher skills $n > n_w$ all look for work but may be involuntarily unemployed (if $\bar{s} < 1$). This is in contrast to Saez (2002), who assumes that agents differ not only in skill levels but also in preferences for full-time leisure. In his setting, all unemployment is voluntary – with unemployed agents exhibiting a higher preference for full-time leisure than the employed agents of the same skill do. Hence, voluntary unemployment is related to not only skill levels but also preferences. In both Saez (2002) and our model, higher skilled agents may be unemployed, but for different reasons. Whereas in Saez (2002) the higher skilled agents without work value leisure highly and thus do not look for work, in our model these agents actively search for work but have been unfortunate enough not to have found a suitable job.

The social planner's problem can now be formulated as

$$\max_{\substack{n_w, u(.), z(.), \\ \omega(.) \geq 0}} \bar{F}(n_w) v(b) + \left[1 - \bar{F}(n_w)\right] \left(-\gamma \bar{s} + (1 - \bar{s}) v(b)\right) \qquad (7)$$

$$+ \int_{n_w}^{n_1} \left\{ \begin{array}{c} \bar{s} u(n) \phi(n) f(n) - \lambda_u(n) \left[u'(n) - \frac{z(n)}{n^2}\right] \\ -\lambda_z(n) \left[z'(n) - \omega(n)\right] + \lambda_E \left[f(n) \bar{s} T(n)\right] + \delta(n) z(n) \end{array} \right\} dn$$

$$-\lambda_E \left\{ b \left[F(n_w) + (1 - F(n_w))(1 - \bar{s})\right] + E \right\}$$

$$-\eta_w \left(\gamma - u(n_w) + v(b)\right),$$

where $\bar{F}(n_w) \equiv \int_{n_0}^{n_w} \phi(n) f(n) dn$. η_w denotes the Lagrange multiplier on the participation constraint for type n_w. It measures the social value of increasing employment by forcing more people to search, and can therefore be interpreted as the value of a work test (and the required information on search intensity) inducing more skills to look for work. The financial gain to the government of inducing type n_w to search for work equals $\bar{s}\left[T(n_w) + b\right]$ and $\left[T(n_w) + b\right]$ equals the implicit tax on search for type n_w.

The shadow values $\lambda_z(n), \delta(n), \eta_w$ are associated with three types of bunching. $\lambda_z(n) < 0$ (implying $\omega(n) \equiv z'(n) = 0$) corresponds to the case

10

in which $z(n)$ and $x(n)$ are constant over a range of skills. We call this *bunching due to violation of monotonicity*. Also the case $\delta(n) > 0$ implies that gross and net incomes are constant over a range of skills. In contrast to bunching on account of violation of monotonicity, however, gross incomes $z(n)$ are necessarily zero over this range, so that utility is constant over the bunching interval (see (3) with $z(n) = 0$). This is called $z = 0$ *bunching*. The search margin gives rise to an additional type of bunching, namely the case in which $n_w > n_0$ and $\eta_w \neq 0$. Accordingly, types $n_0 \leq n < n_w$ do not search. Hence, as voluntarily unemployed, they are bunched without any labor income and collect social assistance b. This we call $s = 0$ *bunching*. Just as $z = 0$ bunching, $s = 0$ bunching can occur only at the bottom of the skill distribution.

This paper focuses on the third type of bunching because three reasons make the first two types of bunching unconvincing formalizations of the bottom of the labor market.[8] First, as mentioned in the introduction, empirical research strongly suggests that the extensive labor supply margin is important, especially for unskilled labor. Second, in the case of bunching due to violation of monotonicity, government spending E and welfare benefits b affect neither marginal tax rates nor unemployment (see Boone and Bovenberg (2002)), which is unrealistic. Third, in the case of $z = 0$ bunching, welfare benefits do impact unemployment and marginal tax rates but in a counter-intuitive way: a higher benefit level lowers unemployment (see Boone and Bovenberg (2002)).

4 Unemployment on the extensive margin

This section explores the case in which the extensive labor supply margin is binding (i.e. $u(n_w) = v(b) + \gamma$ and $\eta_w \neq 0$).[9] Sub-section 4.1 first considers the normal case in which the implicit tax rate on search at the bottom of the labor market is positive $(T(n_w) + b > 0)$, after which sub-section 4.2 investigates how higher public spending and welfare benefits affect optimal tax policy in this case. Sub-section 4.3, finally, shows that with a binding

---

[8]This is true even though the first two types of bunching allow for positive marginal tax rates at the bottom of the income distribution. With $z = 0$ bunching, the unskilled do not work positive hours. This may be interpreted as unemployment.

[9]This section assumes that the monotonicity constraint is met. With a binding search margin, violation of the monotonicity constraint is less likely. The reason is that the search margin takes those agents out of the labor market who, in the absence of a search margin, would be bunched on account of the violation of the monotonicity constraint. With positive search costs and a binding search constraint, $z = 0$ bunching cannot occur (see Boone and Bovenberg (2002), Lemma 3).

participation margin the lowest participating type may actually be subsidized on work $(T(n_w) + b < 0)$.

4.1 Solution optimal tax problem

With an interior solution for n_w, the government finds it optimal not to encourage the least able agents to search for a job; these agents thus remain voluntarily unemployed. Types $n_0 \leq n < n_w$ are bunched without any labor income and instead collect the welfare benefit b. Consequently, the government employs passive welfare benefits instead of active labor-market policies to fight the poverty of these low-skilled agents. The lower bound on the observed wage distribution n_w can be viewed as the effective minimum wage (or wage floor) implied by the welfare benefit b and the taxes on labor income.

With an interior solution for n_w, we find a strikingly simple expression determining the marginal worker

$$b + T(n_w) = \tau(n_w)z(n_w). \tag{8}$$

where $\tau(n_w)$ is the marginal tax rate faced by type n_w, or equivalently $\tau(n_w)$ represents the tax distortion on the intensive (hours) margin. The equation says that for the marginal worker (with skill n_w) the distortion on the extensive margin (i.e. the implicit marginal tax rate between non-employment and employment on the left-hand side of (8)) should equal the distortion on the intensive margin (i.e. the right-hand side of (8)).[10] To understand this equation, one should first note that the marginal type n_w is indifferent between participating and not participating in the labor market so that encouraging this type to look for a job yields no first-order welfare effects for this type. Hence, this agent's utility does not enter the equation determining n_w, and external effects of changes in employment on the government budget determine the optimal employment level. The left-hand side of (8) represents the direct budgetary implications of raising employment by reducing n_w: by bringing a marginal worker into work, the government saves the welfare benefit b and collects additional tax revenue $T(n_w)$. The budgetary implications of behavioral changes of other workers are captured by the right-hand side of (8)). Bringing a marginal type n_w into work encourages workers who are marginally more skilled to work less hours – as they can now more easily

---

[10]This expression can be stated as $\frac{b+T(n_w)}{z(n_w)} = \tau(n_w)$, where the left-hand side is closely related to the replacement rate (in after-tax terms). $\frac{b+T(n_w)}{z(n_w)}$ indicates which part of gross labor income in a marginal job $z(n_w)$ is taxed away through withdrawn welfare benefits b and additional labor taxes $T(n_w)$.

mimic type n_w.[11] The optimal tax system balances the budgetary implications of this latter behavioral response on the intensive margin of the more productive workers (represented by the right-hand side of (8)) with the budgetary implications of the behavioral response on the extensive margin of the marginal workers (represented by the left-hand side of (8)). The government thus faces a trade-off between obtaining revenues from either inducing more agents to search or encouraging a smaller group of agents to work more hours. One can state this dilemma also as one between raising production through more employment or raising it through a higher productivity level (interpreting y as work effort rather than hours worked), or, alternatively, as a trade-off between increasing labor supply on the extensive margin and raising it on the intensive margin.

We characterize the solution to the optimal tax problem in two parts. First, we explain what determines the marginal type n_w and the marginal cost of public funds λ_E. Proposition 3 derives the expressions for the marginal tax rate, utility and production levels for each type n and overall welfare.

Lemma 1 *With voluntary unemployment (i.e. $n_w > n_0$), the solution can be characterized as follows*

$$v'(x_w) = \frac{v(x_w) - v(b) - \gamma}{x_w - b}, \tag{9}$$

$$\lambda_E = \frac{G(n_1) - G(n_w)}{1 - F(n_w) - n_w f(n_w) + \frac{f(n_w)}{v'(x_w)}}, \tag{10}$$

$$n_w(1 - F(n_w))(v(b) + \gamma) + F(n_w)b + \bar{E} \tag{11}$$
$$= \int_{n_w}^{n_1} \{[tf(t) - [1 - F(t)]]v(x(t)) - x(t)f(t)\} dt$$

in the unknowns x_w, n_w and λ_E, where $G(n) \equiv \int_{n_0}^{n} \frac{\phi(t)f(t)}{t}dt$, $\bar{E} \equiv b\frac{1-\bar{s}}{\bar{s}} + \frac{E}{\bar{s}}$, and $x(n)$ is determined by

$$v'(x(n)) = \frac{f(n)}{\frac{G(n_1) - G(n)}{\lambda_E} + f(n)n - [1 - F(n)]}, \tag{12}$$

for $n \geq n_w$ and $x_w = x(n_w)$.
The solution satisfies the properties $x_w > b$, $n_w \in \langle n_0, n_1 \rangle$, $\lambda_E > 0$, and (if

---

[11]If the government could observe skills, it would eliminate the search distortion of the marginal worker (i.e. $b + T(n_w) = 0$) because it does not need to worry about the implications of a zero implicit tax rate between work and non-work for the intensive margin of more efficient workers (see Boone and Bovenberg (2001)).

$\gamma > 0)$

$$\left.\frac{d\lambda_E}{dn_w}\right|_{(n_w,\lambda_E) \; satisfying \; (10)} > 0 > \left.\frac{d\lambda_E}{dn_w}\right|_{(n_w,\lambda_E) \; statisfying \; (11)} . \qquad (13)$$

The system is recursive in that equation (9) determining the consumption level of the marginal worker, $x_w = x(n_w)$, depends on neither λ_E nor n_w. As shown in the appendix, this expression is equivalent to (8). It determines the minimum production level (and given the binding participation constraint also the minimum consumption level) that makes it worthwhile for the government to encourage an agent to search for work and thus employ active labor-market policies rather than welfare benefits as an instrument to alleviate poverty. Production in a marginal job needs to be sufficient to offset the search costs and the implied decline in production of more efficient workers. The consumption level x_w can thus be viewed as the reservation income for the government to have agents search.

To see what determines the consumption gap between the lowest working type and the unemployed, $x_w - b$, we substitute the Taylor expansion $v(b) = v(x_w) + v'(x_w)(b - x_w) + 1/2v''(\xi)(b - x_w)^2$ (where $b < \xi < x_w$) into (9). This yields $\gamma = -1/2v''(\xi)(b - x_w)^2$, so that reservation consumption is a mark-up on the welfare benefit

$$x_w = b + \sqrt{2\gamma/(-v''(\xi))}. \qquad (14)$$

The optimal gap between the consumption level of marginal workers, x_w, and that of the unemployed, b, thus depends on search costs γ and the concavity of utility from consumption, $v(.)$.[12] Large search costs imply that marginal workers need to be rather productive (and thus enjoy relatively high consumption levels) in order to make it worthwhile to have them search for a job. A concave utility function, in contrast, implies that unequal consumption levels of x and b become rather costly. Hence, the government optimally contains the gap between consumption of workers and the unemployed.

We find the optimum by deriving two relationships between λ_E and the marginal skill n_w. In (n_w, λ_E) space (as illustrated in figure 1), labor supply (10), denoted by LS, is upward sloping[13] and the government budget

---

[12]By reducing x_w, a more concave utility function shifts up equation (10) in (n_w, λ_E) space, thereby raising λ_E and hence (from (15)) marginal tax rates. This result contrasts with bunching due to violation of monotonicity and $z = 0$ bunching in which λ_E and marginal tax rates do not depend on the concavity of the utility function. See Boone and Bovenberg (2002) for details.

[13]If one assumes that skills n are lognormally distributed, it is typically the case (as in figure 1) that equation (10) is downward sloping for values of n close to n_0. If $n_w < \underline{n}$

14

constraint (11), denoted GBC, is downward sloping in equilibrium. The LS curve models the optimal trade-off between the incentive constraints (i.e. first-order incentive compatibility and the participation constraint) at different employment levels. In particular, if λ_E increases, consumption as determined by (12) falls for each working type n. Hence, an agent needs to be more productive to meet the participation constraint $(u(n_w) = v(b) + \gamma)$ and n_w goes up (i.e. employment falls). More generous welfare benefits and larger search costs reduce labor supply by raising productivity standards, thereby shifting the LS curve to the right in (n_w, λ_E) space. An alternative interpretation of (10) is that this relationship captures the marginal benefits of using active labor-market policies (rather than passive welfare benefits) as an instrument to combat poverty. These benefits rise with the level of unemployment. More generous welfare benefits and larger search costs reduce these benefits, thereby shifting the upward-sloping relationship (10) downward in (n_w, λ_E) space. The GBC curve represents the costs of these active labor-market policies.

As shown in figure 1, the LS curve features an asymptote at $n_w = \tilde{n}_w$ in (λ_E, n_w) space. In order to interpret $\tilde{n}_w$, we define $g(n_w, \lambda_E)$ as government expenditure that can be financed (with a marginal utility cost of public funds λ_E and a marginal skill level n_w)

$$
\begin{aligned}
g(n_w, \lambda_E) \;\equiv\; & \bar{s} \int_{n_w}^{n_1} \{[tf(t) - [1 - F(t)]] v(x(t)) - f(t) x(t)\} \, dt \\
& - n_w (1 - F(n_w)) \bar{s} (v(b) + \gamma) + b[1 - (1 - F(n_w)) \bar{s}],
\end{aligned}
$$

where $x(t)$ is determined by equation (12).

Lemma 2 *Type $\tilde{n}_w$ provides a zero net contribution to the budget in the sense that*

$$
\left. \frac{\partial g(n_w, \lambda_E)}{\partial n_w} \right|_{(n_w, \lambda_E) \; satisfying \; (10)} = \begin{cases} > 0 & if \quad n_w \in [n_0, \tilde{n}_w\rangle \\ = 0 & if \quad\quad n_w = \tilde{n}_w \\ < 0 & if \quad n_w \in \langle\tilde{n}_w, n_1] \end{cases} .
$$

Furthermore,

$$
T(\tilde{n}_w) + b > 0.
$$

then one gets a violation of the monotonicity constraint and (10) needs to be ironed out to find the equilibrium values of n_w and λ_E (see Boone and Bovenberg (2002) for details). Assuming that the monotonicity constraint is not binding is equivalent to assuming that $n_w > \underline{n}$. Note that for $n_w > \underline{n}$ close to $\underline{n}$ it is the case that small shifts in GBC have little effect on λ_E. This we use in the proof of lemma 5.

15

At $n_w = \tilde{n}_w$, the government collects the largest possible tax revenues (net of spending on welfare benefits). Accordingly, $g(\tilde{n}_w, \lambda_E)$ is the maximum amount of government spending E that can financed. A Leviathan government, which maximizes tax revenues $g(n_w, \lambda_E)$ and thus sets $n_w = \tilde{n}_w$, distorts the extensive margin (i.e. $T(\tilde{n}_w) + b > 0$) because this allows it to collect more tax revenues from types $n_w > \tilde{n}_w$. Indeed, by leaving type n'_w for which $T(n'_w) + b = 0$ out of the labor force, types $n_w > n'_w$ can no longer mimic n'_w, so that the government can raise more tax revenues from these types $n_w > n'_w$. To offset the adverse effects of additional employment on tax revenues from higher types, the government budget must benefit directly from the employment of a marginal type (i.e. $T(\tilde{n}_w) + b > 0$).

In the solution to the planner's problem, we have $n_w < \tilde{n}_w$. Accordingly, employment is subsidized in the sense that a reduction in employment would yield additional revenues (i.e. $\frac{\partial g(n_w, \lambda_E)}{\partial n_w} > 0$). The government budget constraint (11) is thus downward sloping over the relevant range (i.e. around the equilibrium value determined by (10) and (11)):[14] a higher level of employment (i.e. a decrease in n_w) harms the government budget and thus requires an increase in the marginal cost of public funds λ_E to bring the budget back into balance.

Weymark (1987) explores how the rank-order weights $\phi(.)$ affect agents' utilities in a model with quasi linear preferences but without a participation margin. This analysis suggests the following result. Compare two rank order weights $\phi(.)$ and $\tilde{\phi}(.)$ such that $\tilde{\phi}(n) > \phi(n)$ for all $n \in [n_0, \tilde{n}\rangle$ while $\tilde{\phi}(n) < \phi(n)$ for all $n \in \langle\tilde{n}, n_1]$ (and $\int_{n_0}^{n_1} \phi(t) f(t)\, dt = \int_{n_0}^{n_1} \tilde{\phi}(t) f(t)\, dt = 1$). Then under $\tilde{\phi}(.)$ utility is higher for low types, while it is lower for high types compared to utility under $\phi(.)$. Doing the same exercise with a binding participation margin, we find the following additional effects. First, equation (9) implies that x_w is the same under $\phi(.)$ and $\tilde{\phi}(.)$. Second, the LS curve rotates around some skill $n < \tilde{n}$. For values of n_w above $\tilde{n}$ the LS curve rotates downward because $\tilde{G}(n_1) - \tilde{G}(n_w) = \int_{n_w}^{n_1} \frac{f(t)\tilde{\phi}(t)}{t}\,dt < \int_{n_w}^{n_1} \frac{f(t)\phi(t)}{t}\,dt = G(n_1) - G(n_w)$. For values of n_w close to n_0, in contrast, the LS curve rotates upward because $\tilde{G}(n_1) - \tilde{G}(n_w) = \int_{n_w}^{n_1} \frac{f(t)\tilde{\phi}(t)}{t}\,dt > \int_{n_w}^{n_1} \frac{f(t)\phi(t)}{t}\,dt = G(n_1) - G(n_w)$.[15] Thus, if the LS curve rotates around a point above n_w (i.e. $\tilde{\phi}$ puts relatively more weight not only on unemployed but also low

---

[14]This assumes that search costs γ are positive. If $\gamma = 0$, the government budget constraint is horizontal at the equilibrium. Intuitively, in the absence of search costs, inducing agents to search through active labor-market policies does not involve any budgetary costs.

[15]To see this, observe that $\int_{n_0}^{n_1} \left[\tilde{\phi}(t) - \phi(t)\right] f(t)\, dt = 0$ implies that $\int_{n_0}^{n_1} \frac{\tilde{\phi}(t) - \phi(t)}{t} f(t)\, dt > 0$ since $\tilde{\phi}$ exceeds ϕ for high values of $\frac{1}{t}$, while ϕ exceeds $\tilde{\phi}$ for low values of $\frac{1}{t}$.

skilled workers), employment rises as a result of lower average taxes $T(n)$ for n close to n_w. Finally, consider the GBC. The shift in weights from ϕ to $\tilde{\phi}$ reduces consumption for high types (see equation (12) for given λ_E and using the impact of shifting rank-order weights on $G(n_1) - G(n_w)$ derived above) and raises consumption for low types. With the possible exception of types n_w very close to n_0, this shifts the GBC downwards. As a result of the shifts in the two curves, a change in the welfare weights from ϕ to $\tilde{\phi}$ boosts employment. Intuitively, the government draws more low-skilled workers into the labor force to make them better-off (recall that b is fixed here). Furthermore, the lower implicit tax burden on the extensive margin goes together with reduced consumption for high types $n < n_1$ because of higher marginal tax rates on the intensive margin.

We now characterize the marginal tax rate, utility, gross income and welfare.

Proposition 3 *The solution for the marginal tax rate can be written as*

$$\tau(n) = 1 - \frac{\frac{G(n_1) - G(n)}{\lambda_E} + f(n)n - [1 - F(n)]}{n f(n)}, \tag{15}$$

for $n \geq n_w$.
The (in work) utility of a type $n \geq n_w$ is given by

$$u(n) = \frac{1}{n}\left(K + \int_{n_w}^{n} v(x(t))\, dt\right),$$

where K is defined as

$$K \equiv \frac{1}{1 - F(n_w)}\left[\begin{array}{c}\int_{n_w}^{n_1}\{[tf(t) - [1 - F(t)]]\, v(x(t)) - x(t)f(t)\}\, dt \\ -F(n_w)b - \bar{E}\end{array}\right].$$

Gross income of a type $n \geq n_w$ is determined by

$$z(n) = n(v(x(n)) - u(n)).$$

Overall welfare amounts to

$$\begin{aligned}W &= v(b) + \bar{s}(v(b) + \gamma)[n_w(G(n_1) - G(n_w)) - (1 - F(n_w))] + \\ &\quad + \bar{s}\int_{n_w}^{n_1}(G(n_1) - G(n))\, v(x(n))\, dn.\end{aligned}$$

Clearly, the marginal tax rate has the well known property of no distortion at the top (i.e. $\tau(n_1) = 0$). Using equation (11) above to assess the

17

expression for K, one finds that $u(n_w) = \frac{K}{n_w} = v(b) + \gamma$. Since $x(n)$ has been characterized in equation (12) and $u(n)$ is described in terms of $x(n)$ and K, we can employ $u(n) = v(x(n)) - \frac{z(n)}{n}$ to determine gross income $z(n)$. Finally, we can characterize overall welfare in the optimum by a surprisingly simple expression.

Lemma 1 assumes an interior solution for $n_w > n_0$. However, in terms of figure 1, the GBC curve may be everywhere below the LS curve so that $n_w = n_0$. Hence, all agents search for employment and all unemployment is involuntary. The participation margin may then still be binding (in the sense that the unconstrained solution would yield $u(n_0) < v(b) + \gamma$; see Boone and Bovenberg (2002)). In that case, the marginal tax rate on the lowest skill is positive (i.e. $\tau(n_0) > 0$). This result contrasts with the familiar result from the optimal tax literature that, in the absence of bunching at the bottom and a participation constraint, the lowest skill should face a zero marginal tax rate (see Seade (1977)). Intuitively, with a binding search constraint, the government is forced to raise utility $u(n_0) = v(b) + \gamma$ of the least skilled type n_0 above the level (associated with a zero marginal tax rate $\tau(n_0) = 0$) chosen in the absence of this constraint. To prevent higher skills from mimicing the more attractive income bundle of the lowest type n_0, the government distorts work effort of this latter type. In particular, by reducing (net and gross) incomes $x(n_0)$ and $z(n_0)$, a strictly positive marginal tax rate $\tau(n_0) > 0$ makes the income bundle $(x(n_0), z(n_0))$ less attractive for higher types.

This result of a positive marginal tax rate at the bottom parallels that of a monopolist engaging in second-order price discrimination by offering a menu of goods with different qualities and prices. As shown by Mussa and Rosen (1978), such a monopolist offers an inefficiently low quality level to the lowest type in order to be able to extract more rents from buyers featuring higher reservation prices. Also in that application, the participation constraint of the lowest type gives rise to a distortion at the bottom. In our application, the government optimally adjusts the taxation of low-skilled agents in order to induce these agents to continue to look for work, while at the same time minimizing the tax revenues that have to be given up to higher types. This result of a positive marginal tax rate at the bottom illustrates the importance of examining the tax system in conjunction with social assistance. Indeed, the government reduces the tax burden on low-skilled workers as an instrument to contain spending on welfare.

4.2   Comparative statics

This sub-section characterizes the effects of changes in public expenditure E and welfare benefits b if the participation margin is binding. The effects of

E on n_w and λ_E follow directly from figure 1.

Lemma 4 *With $u(n_w) = v(b) + \gamma$ and $n_w > n_0$, we have*

$$\frac{dn_w}{dE} > 0,$$

$$\frac{d\lambda_E}{dE} > 0,$$

$$\frac{dT(n_1)}{dE} > 0.$$

For agents $n > n_w$, we find that

$$\frac{d\tau(n)}{dE} > 0.$$

More public spending raises the marginal costs of active labor-market policies, which as indicated above are a net drain on the government budget (since $n_w < \tilde{n}_w$). As a direct consequence, marginal skills n_w leave the labor market. In the Ebert (1992) version of the model (that is, in the absence of an extensive margin), higher government spending does not impact the shadow price of government spending λ_E and hence leaves marginal tax rates (which given λ_E are determined by (15)) unaffected. The reason why public spending does not affect λ_E is that a uniform increase in labor taxes (i.e. $dT(n) = dT > 0$ for $n_w \geq n_0$) acts as a lump-sum tax in the Ebert model. With the additional behavioral margin of extensive labor supply, however, the government can no longer finance additional government spending in this non-distortionary manner: the binding participation margin implies that a uniform increase in labor taxes causes the marginal worker n_w to stop searching. To alleviate these distortions on the extensive labor-supply margin of low-skilled workers, the government raises a relatively large share of the required additional tax revenues from higher skills. The implied increase in marginal tax rates induces workers to substitute away from consumption towards leisure. As public spending increases, the implied distortions on the intensive labor-supply margin worsen, thereby further increasing the marginal costs of public spending. Equation (34) in the appendix writes marginal cost of public funds (10) as

$$\lambda_E = \left[\frac{G(n_1) - G(n_w)}{1 - F(n_w)} \right] \frac{1}{1 - \chi}, \tag{16}$$

where $\chi \equiv \frac{n_w f(n_w)}{1 - F(n_w)} \frac{b + T(n_w)}{z(n_w)} = \frac{n_w f(n_w)}{1 - F(n_w)} \tau(n_w)$. Whereas the first term (between square brackets) stands for the marginal cost in the absence of behavioral

19

responses, the second term $\frac{1}{1-\chi}$ represents additional costs as a consequence of behavioral responses on the extensive margin.

We now turn to the question whether higher welfare benefits lead to the introduction (or extension) of an EITC. For instance, a government more concerned with the well being of those people without a job tends to raise the benefit b. Does such a government mitigate the adverse effects of higher welfare benefits on labor-market search by reducing the taxes paid by workers at the bottom of the income distribution and phasing out these tax cuts for low-skilled workers by raising marginal tax rates for all working agents?

Lemma 5 *With $u(n_w) = v(b) + \gamma$ and $n_w > n_0$, we have*

$$\frac{dn_w}{db} > 0,$$

and values $\bar{E} < g(\tilde{n}_w, +\infty)$ and $\bar{\gamma} > 0$ exist such that for all $n \in \langle n_w, n_1 \rangle$

$$\frac{d\tau(n)}{db} > 0$$

for $E \in \langle \bar{E}, g(\tilde{n}_w, \lambda_E) \rangle$ and $\gamma \in \langle 0, \bar{\gamma} \rangle$.
If $\tau(n_w) > 0$ and $\frac{d\tau(n)}{dn} \leq 0$ around n_w and a rise in b is not Pareto improving, the results above hold irrespective of the values for $\bar{E}$ and γ. Moreover, in addition to the effect on $\tau(n)$ above, we then have

$$\frac{dT(n)}{db} < 0$$

for n close to n_w.

Larger welfare benefits raise unemployment through two channels. First, the marginal cost of active labor-market policies increase as higher in-work benefits have to be paid to low-skilled agents in order to induce them to enter the labor market. This effect shifts the GBC curve upward. Second, the marginal benefits of active labor-market policies decrease as the welfare system takes on a larger role in combatting poverty. As a direct consequence, less redistribution has to be carried out through the tax system. Indeed, a higher welfare benefit raises the productivity requirements for marginal workers, x_w, so that the LS curve shifts down. With the marginal costs of these policies rising and the associated benefits falling, voluntary unemployment unambiguously rises. With higher welfare benefits, society relies more on these passive benefits rather than on in-work tax benefits to redistribute resources towards low-skilled agents.

The impact of larger welfare benefits on λ_E and hence (see (15)) on marginal tax rates for workers $\tau(n)$ depends on which effect is stronger: the fall in the social benefits of active labor market policies or the increase in costs of these policies. The cost effect dominates if in (n_w, λ_E) space the downward-sloping GBC curve is relatively flat compared to the upward-sloping LS curve. This is in fact the normal case, implying that a higher welfare benefit typically increases the marginal costs of public funds λ_E and hence marginal tax rates.[16] In particular, the cost effect dominates and the normal case thus holds if search costs γ are relatively small (so that the GBC locus is relatively flat) or public spending E is large (so that the LS curve is close to its asymptote $\tilde{n}_w$ and this curve is therefore steep). Furthermore, λ_E and thus marginal tax rates $\tau(n)$ increase with b if marginal tax rates are positive and decline with skill close to the marginal worker n_w and if higher welfare benefits are not Pareto improving.[17] In the normal case, in addition to the unemployed, also marginal workers (i.e. workers with skills close to n_w) benefit from higher welfare benefits as the government reduces the labor taxes paid by these marginal workers. In-work benefits for low-skilled agents and welfare benefits are thus complements: in-work benefits help to offset the impact of more generous social assistance on the participation constraint. These additional in-work tax benefits for low-skilled workers are financed through higher tax rates on skilled workers, implying higher marginal tax rates for workers and a more progressive labor tax system.

The relationship between welfare benefits and the level of in-work benefits appears to be U-shaped. At low welfare benefits, the participation margin is not binding. As welfare benefits are raised from such low initial levels, these benefits absorb the budgetary room for generous in-work benefits as an instrument to fight poverty.[18] At low welfare benefits, social assistance and in-work benefits are thus substitutes in fighting poverty. As welfare benefits are increased further, however, the participation constraint for marginal workers becomes binding and the government needs to raise in-work benefits to encourage agents to look for work. At high levels of social assistance,

---

[16] The benefit effect dominates the cost effect (and the marginal cost of public funds λ_E thus declines) only if the curve (10) is flat enough. This curve is flat if agents are almost bunched in the sense that consumption levels do not rise quickly with n. That is, the equilibrium value of n_w is close to $\underline{n}$ in figure 1. In that case, the higher productivity requirements associated with a higher welfare benefit induce a substantial number of agents to stop searching, thereby saving substantial search costs (especially because the steep government budget curve signals that search costs γ are high).

[17] Although we do not analyze the case in which b is set to maximize welfare, it seems reasonable to exclude the case in which a rise in b yields a Pareto improvement.

[18] For a formal analysis of the comparative statics of higher welfare benefits in a model without a binding participation constraint, see Boone and Bovenberg (2002).

therefore, in-work benefits and welfare benefits become complements. This U-shaped relationship between in-work benefits and social assistance reveals that generous in-work benefits are called for in both countries with low and high welfare benefits, but for different reasons. In countries with low welfare benefits (such as the United States), in-work benefits help to alleviate poverty. In countries with more generous social assistance (such as most European countries), in contrast, in-work benefits protect the incentives to participate in the labor market.

4.3 Negative marginal tax rates

This sub-section considers the case in which in-work benefits are so large that the implicit tax rate on employment is actually negative so that search is subsidized, i.e. $b + T(n_w) < 0$.

Proposition 6 *If $n_w > n_0$, it is possible that*

$$
\begin{aligned}
\eta_w &< 0, \\
b + T(n_w) &< 0, \\
\tau(n_w) &< 0.
\end{aligned}
$$

In this case we have

$$
\frac{dT(n)}{dE} > 0
$$

for all types $n > n_w$ where $\tau(n) < 0$.

With voluntary unemployment, the government can thus optimally provide a search subsidy to marginal workers n_w. The condition for the optimal level of voluntary unemployment $f(n_w)\bar{s}\lambda_E(b + T(n_w)) = \eta_w u'(n_w)$ (see the Appendix for a derivation) implies that a search subsidy for the marginal worker $b + T(n_w) < 0$ is associated with a negative Lagrange multiplier on the participation constraint η_w in the government's optimization problem (7). A further increase in employment thus harms welfare. Moreover, the negative shadow value η_w implies that the participation restriction $u(n_w) = v(b) + \gamma$ is binding from below: $u(n) \leq v(b) + \gamma$ for $n < n_w$ is the relevant constraint. Instead of attracting types $n \geq n_w$ into work, the government wants to keep types $n < n_w$ out of the labor force. This provides a rationale for hiring and firing costs. Employment could also be discouraged by imposing a minimum earnings requirement on jobs. Hence, by prohibiting jobs that earn less than $z_{\min} > 0$, the government facilitates generous tax benefits to low-skilled workers.

In the presence of search subsidies $b + T(n_w) < 0$, the government faces a trade-off between redistributing to agents with low productivity and containing search costs γ.[19] With exogenous welfare benefits, it can redistribute to low productivity agents only by providing rather generous in-work benefits. Additional redistribution, however, results in more entry into the labor force. With a negative implicit tax rate on employment $b + T(n_w) < 0$, the government budget bears part of the search costs associated with this additional entry: an increase in employment implies a direct burden on the government budget, as the additional in-work benefits $-T(n_w)$ exceed the welfare benefit b the government saves. These additional budgetary costs of redistribution stop the government from redistributing more to the bottom of the skill distribution and explain why η_w is negative. Search subsidies are likely to be optimal if not only b but also E is rather low (so that the government has the budgetary means to provide generous in-work benefits to unskilled workers), while at the same time search costs γ are high (so that the government wants to contain entry costs).

Proposition 6 overturns the standard result from the literature on optimal non-linear income taxation that marginal tax rates should be positive in the interior of the skill distribution. The negative marginal tax rate $\tau(n_w)$[20] can be explained in two ways. First, the government likes to redistribute to less productive agents but at the same time wants to avoid excessive search. Search is excessive here due to the externality on the government budget constraint: agents at the bottom of the skill distribution do not take into account that finding a job is a drain on the government budget $T(n_w) + b < 0$. To reconcile the objectives of alleviating poverty and avoiding

---

[19]In the presence of search subsidies, voluntary unemployment can exist only with positive search costs. Indeed, agents refrain from search only if they have to incur search costs in order to take advantage of the higher public benefits in employment (i.e. $-T(n_w) > b$).

[20]Crucial for the result that marginal tax rates can be negative is that the government can not freely set b to optimize social welfare in our setting. If the government could increase the welfare benefit b, it would employ this instrument to alleviate poverty. In this way, it could redistribute resources to low-skilled agents without having to rely on in-work benefits that require low-skilled agents to engage in costly search. With inward migration, however, the government may find it hard to affect the outside option of migrants. Indeed, a government facing inward migration of low-skilled labor may want to make the tax system less progressive compared to the case in which it does not face the participation constraint of migrants. Moreover, the welfare benefit b may be set by another government agency than the tax authorities determining labor taxes. The tax authorities may thus have to treat b as being exogenous.

Alternatively, we could allow for a more general search cost function $\gamma(s)$ then the one used above. In that case, we show in Boone and Bovenberg (2002) that we get the same expression for the marginal tax rate as Saez (2002). Saez shows that in that case negative marginal tax rates are possible even if the government sets b optimally as well.

excessive search, it grants the largest in-work subsidies to skill types above n_w rather than to the marginal worker n_w (as this would draw types $n < n_w$ into the labor force). Hence, the marginal tax rate is negative at n_w.

The second way to explain the negative marginal tax rates is as follows. In order to separate the types at the bottom of the labor force (and thus make it unattractive for the types below n_w to mimic type n_w by searching for a job), the government induces marginal workers n_w to supply excessive labor on the intensive margin (which is cheaper for type n_w than for types below n_w), implying that excessive labor supply on the extensive margin spreads to the intensive margin. Indeed, in contrast to traditional optimal tax problems, the self selection constraints are binding from below in the sense that they prevent low types from mimicing high types (rather than preventing high types from mimicing low types).

As noted in lemma 4, more public spending reduces employment. In the presence of search subsidies, this helps to alleviate excessive employment. This side benefit of higher public spending contains the marginal cost of public funds. Indeed, expression (16) implies that the marginal costs of public funds are below $G(n_1)$, its level in the Ebert (1992) model without a binding participation constraint.[21]

5 Conclusions

This paper has explored how the income tax system should optimally respond to an exogenously given welfare benefit in the presence of costly labor-market search and non-verifiable skills. We showed that optimal unemployment is determined by the requirement that distortions on the extensive margin balance those on the intensive margin. On the one hand, generous in-work benefits help to alleviate distortions on the participation margin by encouraging more low-skilled workers to actively look for work. On the other hand, such benefits make it more attractive for high-ability agents to mimic lower ability agents, thereby distorting work effort. The government thus faces a trade-off between boosting (low productive) employment and raising work effort of higher skilled workers. If the government lacks information on individual skills, a distorted participation margin in the form of a positive implicit tax rate between non-employment and employment is therefore the price for combatting poverty while at the same time protecting labor supply

---

[21]The first term at the right-hand side of (16), $\frac{G(n_1)-G(n_w)}{1-F(n_w)}$, is declining in n_w and is thus smaller than $G(n_1)$ if $n_w > n_0$. The second term is smaller than one if $T(n_w)+b < 0$. Hence, $\lambda_E \leq G(n_1)$.

of higher skilled workers. A similar trade-off appears in determining optimal retirement schemes. In particular, rather than linking public retirement benefits to the retirement age in an actuarially fair way, the government may want to favor early retirement to aid low productivity individuals suffering from poor health (see Cremer, Lozachmeur, and Pestieau (2002)).

In the presence of low search costs, low welfare benefits, a concentrated skill distribution and large public spending, the government may find it optimal to employ a progressive tax system providing generous in-work benefits to low-skilled workers in order to induce all agents to search. Such a progressive tax system features a positive marginal tax rate at the bottom. This contrasts with the familiar result from the optimal tax literature that, in the absence of bunching at the bottom, those with the lowest skills should face a zero marginal tax rate (see Seade (1977)). This new result shows that the welfare system and the participation margin may importantly affect the optimal tax system.

Social assistance and positive search costs may overturn also the well-known result from the optimal tax literature that marginal taxes should be positive in the interior of the income distribution. We showed that the government may optimally increase in-work benefits with gross income (implying a negative marginal tax on work effort) in order to limit excessive entry into the labor market. In particular, whereas low tax levels combined with positive marginal tax rates at the bottom help to encourage search if labor-force participation of low-skilled workers is taxed on a net basis, negative marginal taxes rates for low-skilled workers help to discourage excessive entry of low-skilled workers if this entry is subsidized. This latter case provides a rationale for minimum wages and hiring and firing costs.

The incorporation of labor-market imperfections and the welfare system into a model of optimal non-linear income taxation enabled us to investigate how these new elements impact the optimal income tax. We showed, for example, that more generous welfare benefits tend to raise marginal tax rates, as the government cuts the average tax burden on low-skilled workers in order to encourage these workers to continue to look for a job. Accordingly, skilled workers finance not only more generous social assistance, but also a lower tax burden of low-skilled workers.

In future research, we would like to investigate optimal tax policy if the government can simultaneously set welfare benefits, search obligations, and other categorical social insurance benefits (such as disability benefits based on a signal of skill type). A study of these issues would need to allow for imperfect information on search behavior, household structure, and skill types. In exploring the optimal trade-off between passive welfare benefits and active labor-market policies, we also intend to account for negative external effects

of unemployment.

6   References

Atkinson, A.B., 1995, *Public Economics in Action. The Basic Income/Flat Tax Proposal* (Clarendon Press, Oxford).

Blundell, R., 2001, Welfare-to-Work: Which Policies Work and Why?, Keynes Lecture in Economics.

Boadway, R., K. Cuff, and M. Marchand, 2000, Optimal Income Taxation with Quasi-Linear Preferences Revisited, *Journal of Public Economic Theory,* Vol. 2, pp. 435-460.

Boone, J., and A.L. Bovenberg, 2001, 'Unemployment vs. In-work Benefits with Search Unemployment and Observable Abilities,' Center Discussion Paper No. 77, Tilburg University, the Netherlands.

Boone, J., and A.L. Bovenberg, 2002, The optimal taxation of unskilled labour with job search and social assistance, CEPR discussion paper No. 3446.

Chone, P., and G. Laroque, 2001, 'Optimal Incentives for Labor Force Participation,' INSEE Working Paper 2001-25, Paris.

Cremer, H., J. Lozachmeur, and P. Pestieau, 2002, 'Social Security, Retirement Age and Optimal Income Taxation,' CESifo Working Paper, No. 693 (paper presented at 2002 Trans-Atlantic Public Economics Seminar, London).

Diamond, P.A., 1980, 'Income Taxation with Fixed Hours of Work,' *Journal of Public Economics*, Vol. 13, pp. 101-110.

Diamond, P.A., 1998, 'Optimal Income Taxation: An Example with a U-Shaped Pattern of Optimal Marginal Tax Rate,' *American Economic Review*, Vol. 88, pp. 83-95.

Ebert, U., 1992, 'A Reexamination of the Optimal Nonlinear Income Tax,' *Journal of Public Economics*, Vol. 49, pp. 47-73.

Eissa, N., and J. Liebman, 1996, 'Labor Supply Response to the Earned Income Tax Credit', *Quarterly Journal of Economics*, Vol. 111 (2), pp. 605-637.

Fudenberg, D., and J. Tirole, 1991, *Game Theory* (MIT Press, Cambridge).

Heckman, J., 1993, 'What has been Learned about Labor Supply in the Past Twenty Years?', *American Economic Review*, Vol. 83 (2), pp. 116-121.

Kamien, M.I., and N.L. Schwartz, 1981, *Dynamic optimization: The calculus of variations and optimal control in economics and management* (North-Holland, Amsterdam).

Kanbur, R., M. Keen, and M. Tuomala, 'Optimal Non-Linear Income Taxation for the Alleviation of Income-Poverty,' *European Economic Review*, Vol. 38, pp. 1613-1632.

Kimmel, J., and T. Kniesner, 1998, 'New Evidence on Labor Supply: Employment versus Hours Elasticities by Sex and Marital Status', *Journal of Monetary Economics,* Vol. 42, pp. 289-301.

Lollivier, S., and J. Rochet, 1983, Bunching and Second-Order Conditions: A Note on Optimal Tax Theory, *Journal of Economic Theory*, Vol. 31, pp. 392-400.

Meyer, B. and D. Rosenbaum, 2001, Welfare, the Earned Income Tax Credit, and the Labor Supply of Single Mothers, *Quarterly Journal of Economics*, Vol. 66, pp. 1063-1114.

Meyer, B., 2002, 'Labor Supply at the Extensive and Intensive Margins: The EITC, Welfare and Hours Worked', Mimeo.

Mirrlees, J.A., 1971, An Exploration in the Theory of Optimal Income Taxation,' *Review of Economic Studies*, Vol. 38, pp. 175-208.

Mortensen, D.T., and C.A. Pissarides, 1999, New Developments in Models of Search in the Labor Market, in O. Ashenfelter and D. Card (eds.), *Handbook of Labor Economics*, Vol. 3 (North-Holland, Amsterdam).

Mussa, M. and S. Rosen, 1978, Monopoly and product quality, *Journal of Economic Theory*, Vol. 18, pp. 301-317.

OECD, 2001, *Education at a Glance*, Paris.

Rochet, J., and L. Stole, 2002, Nonlinear Pricing with Random Participation, *Review of Economic Studies*, Vol. 69, pp. 277-311.

Sadka, E., 1976, On Income Distribution, Incentive Effects and Optimal Income Taxation, *Review of Economic Studies*, Vol. 43, pp. 261-267.

Saez, E., 2002, Optimal Income Transfer Programs: Intensive Versus Extensive Labor Supply Responses, *Quarterly Journal of Economics* (forthcoming).

Seade, J.K., 1977, On the Shape of Optimal Tax Schedules, *Journal of Public Economics*, Vol. 7, pp. 203-235.

Weymark, J.A., 1986, A Reduced-Form Optimal Nonlinear Income Tax Problem, *Journal of Public Economics*, Vol. 30, pp. 199-217.

Weymark, J.A., 1987, Comparative Static Properties of Optimal Nonlinear Taxes, *Econometrica*, Vol. 55, pp. 1165-1185.

7 Appendix

Proof of Lemma 1

First we state the following result for reference below (for a proof, see Boone and Bovenberg (2002)).

Lemma 7 *With voluntary unemployment (i.e. $n_w > n_0$), the participation margin is strictly binding (i.e. $u(n_w) = \gamma + v(b)$). Furthermore, if $\gamma > 0$, a binding participation constraint $u(n_w) = \gamma + v(b)$ implies $z(n) > 0$ for $n \geq n_w$.*

The Euler equations for maximizing (7) with respect to $\omega(.), u(.)$ and $z(.)$ can be written as

$$\omega(n) = \arg\max_{\omega \geq 0} \lambda_z(n)\omega, \tag{17}$$

$$\lambda'_u(n) = \bar{s}f(n)\left(\frac{\lambda_E}{v'(x(n))} - \phi(n)\right), \tag{18}$$

$$\lambda'_z(n) = -\frac{\lambda_u(n)}{n^2} + \lambda_E f(n)\bar{s}\left(\frac{1}{nv'(x(n))} - 1\right), \tag{19}$$

We assume that the monotonicity conditions are met so that $\lambda_z(n) = 0$ and thus $\lambda'_z(n) = 0$ for all n. Expression (19) can therefore be written as

$$\frac{1}{v'(x(n))} = n + \frac{1}{n}\frac{\lambda_u(n)}{\lambda_E f(n)\bar{s}}. \tag{20}$$

The first-order condition for maximizing individual utility with respect to $z(n)$ in equation (2) amounts to

$$v'\left(z(n) - \tilde{T}(z(n))\right)\left(1 - \tilde{T}'(z(n))\right) - \frac{1}{n} = 0,$$

or equivalently,

$$v'(x(n)) = \frac{1}{n(1-\tau(n))}.$$

Using this in equation (20) to eliminate $v'(x(n))$, we find

$$\tau(n) = \frac{-\lambda_u(n)}{\lambda_E n^2 f(n)\bar{s}}. \tag{21}$$

Substituting equation (20) into (18) to eliminate $v'(x(n))$, we arrive at

$$\lambda'_u(n) = \frac{1}{n}\lambda_u(n) + \lambda_E \bar{s} f(n) n - \bar{s} f(n)\phi(n) \tag{22}$$

This is a linear differential equation that can be solved analytically (using the method of the varying constant):

$$\lambda_u(n) = n\left[c_0 + \bar{s}\left(\lambda_E \int_{n_0}^{n} f(t)\,dt - \int_{n_0}^{n}\frac{f(t)\phi(t)}{t}dt\right)\right] \tag{23}$$

for some constant c_0.

With an interior solution $n_w > n_0$, the first-order condition for maximizing (7) with respect to n_w amounts to

$$\eta_w u' (n_w) = f (n_w) \{\bar{s} [-\gamma - v (b) + u (n_w) + \lambda_E (b + T(n_w))]\} . \qquad (24)$$

The transversality conditions[22] are

$$
\begin{aligned}
\lambda_u (n_w) + \eta_w &= 0, &\qquad (25)\\
\lambda_u (n_1) &= 0, &\qquad (26)\\
\lambda_z (n_w) &= 0, &\qquad (27)\\
\lambda_z (n_1) &= 0. &\qquad (28)
\end{aligned}
$$

The second transversality condition implies that differential equation (22) for $\lambda_u (n)$ can be solved as

$$\lambda_u (n) = -n\bar{s} [\lambda_E (1 - F (n)) - (G (n_1) - G (n))] , \qquad (29)$$

which implies (together with (20)) that (12) holds. Solving this expression for λ_E, we arrive at (10).

Using (25) to eliminate $\lambda_u (n_w)$, we can write the solution (29) for $n = n_w$ as

$$-\frac{\eta_w}{n_w} + \bar{s} [\lambda_E (1 - F (n_w)) - (G (n_1) - G (n_w))] = 0. \qquad (30)$$

We know from lemma 7 that an interior solution $n_w > n_0$ implies $u (n_w) = v (b) + \gamma$, so that we can write (24) as

$$\eta_w = \frac{f (n_w) \bar{s} \lambda_E (b + T (n_w))}{u' (n_w)}. \qquad (31)$$

Substituting (25) into (21) to eliminate $\lambda_u (n_w)$, and using (31) to eliminate η_w from the resulting expression, we obtain

$$
\begin{aligned}
\tau (n_w) &= \frac{b + T (n_w)}{u' (n_w) \, n_w^2}\\
&= \frac{b + T (n_w)}{z (n_w)}, \qquad (32)
\end{aligned}
$$

where we have used the incentive compatibility $u' (n_w) = \frac{z(n_w)}{n_w^2}$ to derive the second expression. Since $v' (x_w) = \frac{1}{n_w(1-\tau(n_w))}$, we find (using (32) to

---

[22] For details see Boone and Bovenberg (2002).

eliminate $\tau(n_w)$)

$$
\begin{aligned}
v'(x_w) &= \frac{1}{n_w} \frac{z(n_w)}{z(n_w) - b - T(n_w)} \\
&= \frac{v(x_w) - \gamma - v(b)}{x_w - b},
\end{aligned} \tag{33}
$$

where we have used (from the binding participation constraint) $z(n_w) = n_w(v(x_w) - u(n_w)) = n_w(v(x_w) - v(b) - \gamma)$ and $T(n_w) = z(n_w) - x_w$ to eliminate $z(n_w)$ and $T(n_w)$ from the first equality in (33).

Substituting (30) into (31) to eliminate η_w, we find

$$
\begin{aligned}
\lambda_E &= \frac{\frac{G(n_1) - G(n_w)}{1 - F(n_w)}}{1 - \frac{f(n_w)}{1 - F(n_w)} \frac{1}{u'(n_w)} \frac{T(n_w) + b}{n_w}} \\
&= \frac{G(n_1) - G(n_w)}{1 - F(n_w)} \frac{1}{1 - \chi},
\end{aligned} \tag{34}
$$

where $\chi \equiv \frac{f(n_w)n_w}{1 - F(n_w)} \frac{T(n_w) + b}{z(n_w)}$ and we have used incentive compatibility $u'(n_w) = \frac{z(n_w)}{n_w^2}$ to eliminate $u'(n_w)$.

The government budget constraint is given by

$$
\int_{n_w}^{n_1} f(n)\bar{s}\left[z(n) - x(n)\right] dn = E + \left[F(n_w) + (1 - F(n_w))(1 - \bar{s})\right]b \tag{35}
$$

and the definition of $u(n)$ allows us to write

$$
z(n) = n(v(x(n)) - u(n)). \tag{36}
$$

To determine what $u(n)$ looks like, we substitute this expression for $z(n)$ into the incentive compatibility constraint $u'(n) = \frac{z(n)}{n^2}$ to arrive at the following differential equation:

$$
u'(n) = -\frac{1}{n}u(n) + \frac{1}{n}v(x(n)).
$$

This linear differential equation can be solved analytically (with the method of varying constant):

$$
u(n) = \frac{1}{n}\left(K_w + \int_{n_w}^{n} v(x(t)) dt\right), \tag{37}
$$

$$
z(n) = nv(x(n)) - K_w - \int_{n_w}^{n} v(x(t)) dt,
$$

for some constant K_w. We solve for this constant by substituting the last equation into (35) to eliminate $z(n)$ and solving for K_w:

$$K_w = \frac{1}{1 - F(n_w)} \left[\begin{array}{c} \int_{n_w}^{n_1} \{ [t v(x(t)) - x(t)] f(t) - [1 - F(t)] v(x(t)) \} dt \\ - b F(n_w) - \bar{E} \end{array} \right].$$

Using (37), $u(n_w) = v(b) + \gamma$ can be written as $\frac{1}{n_w} K_w = v(b) + \gamma$. Substituting the expression for K_w above, we establish (11).

We now show that the solution satisfies (13). First, consider (10). The numerator is positive and decreasing in n_w. Denote the denominator by

$$D(n_w) \equiv 1 - F(n_w) - n_w f(n_w) + \frac{f(n_w)}{v'(x_w)}. \tag{38}$$

The assumptions $1 - n_0 f(n_0) > 0$ and $v'(\hat{x}_w) > \frac{1}{n_1}$ imply

$$
\begin{aligned}
D(n_0) &> 0, \\
D(n_1) &< 0.
\end{aligned}
$$

Let $\tilde{n}_w$ denote the smallest value such that $D(\tilde{n}_w) = 0$ (by continuity of $f(.)$ such a value exists).[23] Hence, in (n_w, λ_E) space, (10) features an asymptote at $n_w = \tilde{n}_w$ (see figure 1). Since the numerator is positive, $\lim_{n_w \uparrow \tilde{n}_w} \lambda_E(n_w) = +\infty$. Furthermore, (10) is monotonically increasing in (n_w, λ_E) space for $n_w \in [n_0, \tilde{n}_w]$ if $x'(n) > 0$ for all n (as we assume here). To see this, suppose that $\left. \frac{\partial \lambda_E}{\partial n_w} \right|_{(10)} > 0$ does not hold. Since equation (10) is identical to equation (12) (where the relation between $v'(.)$ and λ_E is a positive one), this implies that $v'(x(n))$ is increasing in n for some $n > n_w$, so that $x'(n) < 0$ for these values. This contradicts the monotonicity requirement $x'(n), z'(n) \geq 0$. Accordingly, a non-binding monotonicity requirement implies $\left. \frac{\partial \lambda_E}{\partial n_w} \right|_{(10)} > 0$.

To determine the slope of (11), we define

$$
\begin{aligned}
\phi(n_w, \lambda_E) &\equiv \int_{n_w}^{n_1} \{ [t v(x(t, \lambda_E)) - x(t, \lambda_E)] f(t) - [1 - F(t)] v(x(t, \lambda_E)) \} dt \\
&\quad - \{ n_w (1 - F(n_w)) (v(b) + \gamma) + b F(n_w) + \bar{E} \}.
\end{aligned}
$$

To find the slope of (11) (which is $\phi(n_w, \lambda_E) = 0$) in (n_w, λ_E) space, we determine $\phi_{n_w} = \frac{\partial \phi}{\partial n_w}$ and $\phi_{\lambda_E} = \frac{\partial \phi}{\partial \lambda_E}$

$$
\begin{aligned}
\phi_{n_w} &= -[n_w v(x(n_w, \lambda_E)) - x(n_w, \lambda_E)] f(n_w) + [1 - F(n_w)] v(x(n_w, \lambda_E)) \\
&\quad - \{ [1 - F(n_w)] (v(b) + \gamma) - n_w f(n_w) (v(b) + \gamma) + f(n_w) b \} \\
&= -(v(x_w) - v(b) - \gamma)(n_w f(n_w) - [1 - F(n_w)]) + f(n_w)(x_w - b) \\
&= (x_w - b) v'(x_w) D(n_w) > 0, \tag{39}
\end{aligned}
$$

---

[23] If $f(.)$ satisfies the monotone hazard rate property, $\tilde{n}_w$ is the unique solution to $D(\tilde{n}_w) = 0$.

where $D(n_w)$ is defined in (38), and the last equality holds only for the values of λ_E and n_w satisfying (10) (since we used (9) to eliminate $(v(x_w) - v(b) - \gamma)$).

$$
\begin{aligned}
\phi_{\lambda_E} &= \int_{n_w}^{n_1} \left\{ [tv'(x(t, \lambda_E)) - 1] f(t) - [1 - F(t)] v'(x(t, \lambda_E)) \right\} \frac{dx(t, \lambda_E)}{d\lambda_E} dt \\
&= \int_{n_w}^{n_1} \left\{ v'(x(t, \lambda_E)) (tf(t) - [1 - F(t)]) - f(t) \right\} \frac{dx(t, \lambda_E)}{d\lambda_E} dt \\
&= \int_{n_w}^{n_1} f(t) \frac{\frac{G(n_1) - G(t)}{\lambda_E}}{\frac{G(n_1) - G(t)}{\lambda_E} + f(t) t - [1 - F(t)]} \left[-\frac{dx(t, \lambda_E)}{d\lambda_E} \right] dt > 0, \ (40)
\end{aligned}
$$

where we have used equation (12) to derive the third equality by eliminating $v'(x(t, \lambda_E))$. The sign of ϕ_{λ_E} follows because the concavity of $v(x)$ together with (12) imply $\frac{dx(t, \lambda_E)}{d\lambda_E} < 0$. With $\phi_{n_w} > 0$ and $\phi_{\lambda_E} > 0$, (11) is downward sloping in (n_w, λ_E) space at the point where (11) and (10) intersect.

Finally, the expression for W in the proposition can be derived by writing individual utility (from equation (37)) as $u(n) = \frac{1}{n}\left(n_w u(n_w) + \int_{n_w}^n v(x(t)) dt\right)$, substituting this expression into welfare

$$
W = \bar{F}(n_w) v(b) + \left[1 - \bar{F}(n_w)\right] (-\gamma \bar{s} + (1 - \bar{s}) v(b)) + \int_{n_w}^{n_1} \bar{s} f(n) \phi(n) u(n) dn
$$

to eliminate $u(n)$, and employing partial integration. This yields

$$
\begin{aligned}
W &= v(b)(1 - \bar{s}) + \bar{s} F(n_w)(v(b) + \gamma) + \bar{s} u(n_w) n_w (G(n_1) - G(n_w)) + \\
&\quad + \bar{s} \int_{n_w}^{n_1} (G(n_1) - G(n)) v(x(n)) dn.
\end{aligned}
$$

Substitution of $u(n_w) = v(b) + \gamma$ yields the expression in the proposition. Q.E.D.

Proof of lemma 2

From the definition of the asymptote $\tilde{n}_w$ above, we have

$$
D(n_w) \begin{cases} > 0 & \text{if } n_w < \tilde{n}_w \\ = 0 & \text{if } n_w = \tilde{n}_w \\ < 0 & \text{if } n_w > \tilde{n}_w \end{cases}.
$$

The result in the lemma on $\frac{\partial g}{\partial n_w}$ follows from (39), $x_w > b$ and $v'(x_w) > 0$.

From the definition of $\tilde{n}_w$ $(D(\tilde{n}_w) = 0$, i.e. (38)), we observe

$$
f(\tilde{n}_w)\left(\frac{1}{v'(x_w)} - \tilde{n}_w\right) = -(1 - F(\tilde{n}_w)) < 0,
$$

so that $v'(x(\tilde{n}_w)) = v'(x_w) > \frac{1}{\tilde{n}_w}$. This implies $\tau(\tilde{n}_w) > 0$ and thus (from (32)) $T(\tilde{n}_w) + b > 0$. Q.E.D.

Proof of Proposition 3

See proof of Lemma 1 above.

Proof of Lemma 4

Differentiating equations (10) and (11) with respect to n_w, λ_E, E, and b, we find

$$
\begin{pmatrix} \phi_{n_w} & \phi_{\lambda_E} \\ -\left.\frac{d\lambda_E}{dn_w}\right|_{(10)} & 1 \end{pmatrix} \begin{pmatrix} dn_w \\ d\lambda_E \end{pmatrix} = \begin{pmatrix} \frac{1}{s}dE + \begin{bmatrix} n_w(1 - F(n_w))v'(b) \\ +\frac{1}{s} - (1 - F(n_w)) \end{bmatrix} db \\ \frac{d\lambda_E}{dv'(x_w)}\frac{dv'(x_w)}{db}db \end{pmatrix},
$$

where ϕ_{n_w} and ϕ_{λ_E} are defined in equations (39) and (40). The determinant of the matrix on the left-hand side equals

$$
\det = \begin{vmatrix} \phi_{n_w} & \phi_{\lambda_E} \\ -\left.\frac{d\lambda_E}{dn_w}\right|_{(10)} & 1 \end{vmatrix} = \phi_{n_w} + \phi_{\lambda_E}\left.\frac{d\lambda_E}{dn_w}\right|_{(10)} > 0.
$$

Hence,

$$
\begin{pmatrix} dn_w \\ d\lambda_E \end{pmatrix} = \frac{1}{\det} \begin{pmatrix} 1 & -\phi_{\lambda_E} \\ \left.\frac{d\lambda_E}{dn_w}\right|_{(10)} & \phi_{n_w} \end{pmatrix} \times \tag{41}
$$

$$
\begin{pmatrix} \frac{1}{s}dE + \begin{bmatrix} n_w(1 - F(n_w))v'(b) \\ +\frac{1}{s} - (1 - F(n_w)) \end{bmatrix} db \\ \frac{d\lambda_E}{dv'(x_w)}\frac{dv'(x_w)}{db}db \end{pmatrix}.
$$

We thus find

$$
\frac{dn_w}{dE} = \frac{1}{\det}\frac{1}{s} > 0
$$

$$
\frac{d\lambda_E}{dE} = \frac{1}{\det}\left.\frac{d\lambda_E}{dn_w}\right|_{(10)}\frac{1}{s} > 0.
$$

From (15) and (12), the effect of E on λ_E implies that

$$
\frac{d\tau(n)}{dE} > 0,
$$

$$
\frac{dx(n)}{dE} < 0.
$$

Writing $u(n)$ as

$$
u(n) = \frac{n_w}{n}(v(b) + \gamma) + \frac{1}{n}\int_{n_w}^{n} v(x(t))\, dt,
$$

33

we find that

$$
\frac{du\left(n\right)}{dE} = \frac{1}{n}\underbrace{\left[v\left(b\right)+\gamma-v\left(x\left(n_w\right)\right)\right]}_{=-\frac{z\left(n_w\right)}{n_w}<0}\underbrace{\frac{dn_w}{dE}}_{>0}+\frac{1}{n}\int_{n_w}^{n}v'\left(x\left(t\right)\right)\underbrace{\frac{dx\left(t\right)}{dE}}_{<0}dt
$$

$$
< \quad 0.
$$

Writing $T\left(n\right)=z\left(n\right)-x\left(n\right)=n\left(v\left(x\left(n\right)\right)-u\left(n\right)\right)-x\left(n\right),$ we find

$$
\frac{dT\left(n\right)}{dE} = \underbrace{-n\frac{du\left(n\right)}{dE}}_{>0}+\underbrace{\left[nv'\left(x\left(n\right)\right)-1\right]}_{=\frac{\tau\left(n\right)}{1-\tau\left(n\right)}\geq0}\underbrace{\frac{dx\left(n\right)}{dE}}_{<0}. \tag{42}
$$

Since $\tau\left(n_1\right)=0$, we find $\frac{dT\left(n_1\right)}{dE}>0$. Q.E.D.

Proof of Lemma 5

Note that (from (12)) $\frac{d\lambda_E}{dv'\left(x_w\right)}>0$ and (from differentiating (9)) $\frac{dv'\left(x_w\right)}{db}<0$ so that (41) yields

$$
\frac{dn_w}{db}>0.
$$

We also have

$$
\frac{d\lambda_E}{db}=\frac{1}{\det}\left[\begin{array}{c}\left.\frac{d\lambda_E}{dn_w}\right|_{(10)}\underbrace{\left[n_w\left(1-F\left(n_w\right)\right)v'\left(b\right)+\frac{1}{\bar{s}}-\left(1-F\left(n_w\right)\right)\right]}_{>0}+\\[2mm] +\phi_{n_w}\frac{d\lambda_E}{dv'\left(x_w\right)}\underbrace{\frac{dv'\left(x_w\right)}{db}}_{<0}\end{array}\right].
$$

We thus obtain $\frac{d\lambda_E}{db}>0$ if ϕ_{n_w} is close to 0, and $\frac{d\lambda_E}{db}<0$ if $\left.\frac{d\lambda_E}{dn_w}\right|_{(10)}$ is close to 0. We consider these cases in turn. First, (39) implies that ϕ_{n_w} is close to 0 if either x_w is close to b or if $D\left(n_w\right)$ is close to 0. (14) reveals that x_w-b is close to 0 if γ is close to zero. Hence, $\bar{\gamma}>0$ exists such that for all $\gamma\in\langle0,\bar{\gamma}\rangle$ ϕ_{n_w} is small enough to ensure $\frac{d\lambda_E}{db}>0$. Moreover, by continuity, $D\left(n_w\right)$ is close to 0 if n_w is close to the asymptote $\tilde{n}_w$ in figure 1. By raising E, the downward-sloping government budget constraint in figure 1 shifts upward so that we can obtain an equilibrium value arbitrarily close to $\tilde{n}_w$. Therefore, by choosing E sufficiently large (but below maximum government revenues $g\left(\tilde{n}_w,+\infty\right)$), ϕ_{n_w} is close enough to 0 so that $\frac{d\lambda_E}{db}>0$.

If (10) is downward sloping around n_0 (as depicted in figure 1), a value for n_w exists (denoted by $\underline{n}$ in figure 1) where $\left.\frac{d\lambda_E}{dn_w}\right|_{(10)}=0$. In this case, we can find values for E such that the intersection of (11) and (10) is at $n_w>\underline{n}$

34

but is close enough to $\underline{n}$ so that $\frac{d\lambda_E}{db} < 0$. More precisely, let E_* denote the value of E such that the intersection is at $\underline{n}$. Then a value $E^* > E_*$ exists such that $\frac{d\lambda_E}{db} < 0$ for all $E \in \langle E_*, E^* \rangle$.

The effects of b on $x(n)$ and $\tau(n)$ follow immediately from $\frac{d\lambda_E}{db}$ by using (12) and (15), respectively.

Now we argue that $\frac{d\lambda_E}{db} < 0$ can be excluded if $\frac{d\tau(n)}{dn} < 0$ around n_w and if a rise in b does not lead to a Pareto improvement. More precisely, we show that $\frac{d\lambda_E}{db} < 0$ and $\frac{d\tau(n)}{dn} < 0$ around n_w imply that raising b does yield a Pareto improvement.

Utility can be written (from (37)) as

$$u(n) = \frac{1}{n} \left(u(n_w) + \int_{n_w}^n v(x(t)) \, dt \right) = \frac{1}{n} \left(v(b) + \gamma + \int_{n_w}^n v(x(t)) \, dt \right)$$

Taking the derivative with respect to b, we arrive at

$$
\begin{aligned}
\frac{du(n)}{db} &= \left[\begin{array}{c} \frac{1}{n}(v(b) + \gamma - v(x_w)) \frac{dn_w}{db} + \\ \frac{n_w}{n} v'(b) + \frac{1}{n} \int_{n_w}^n v'(x(t)) \frac{dx(t)}{db} dt \end{array} \right] \qquad (43) \\
&= \frac{n_w}{n} \left[v'(b) - \frac{v'(x_w)}{n_w}(x_w - b) \frac{dn_w}{db} \right] + \frac{1}{n} \int_{n_w}^n v'(x(t)) \frac{dx(t)}{db} dt,
\end{aligned}
$$

where we have used (9) to eliminate $(v(b) + \gamma - v(x_w))$. Differentiation of $v'(x_w) = 1/[n_w(1 - \tau(n_w))]$ yields $dn_w/dx_w \leq n_w^2(1 - \tau(n_w))(-v''(x_w))$ if $d\tau(n)/dn \leq 0$ around $n = n_w$. Similarly, differentiating (9), we find

$$\frac{dx_w}{db} = \frac{v'(b) - v'(x_w)}{(x_w - b)(-v''(x_w))}.$$

Accordingly, $\frac{dn_w}{db} = \frac{dn_w}{dx_w} \frac{dx_w}{db} \leq n_w^2 \frac{v'(b) - v'(x_w)}{x_w - b}(1 - \tau(n_w))$. Substitution of this inequality in (43) to eliminate $\frac{dn_w}{db}$, we obtain

$$\frac{du(n)}{db} \geq \frac{n_w}{n} \left[v'(b) - v'(x_w) n_w (1 - \tau(n_w))(v'(b) - v'(x_w)) \right] + \frac{1}{n} \int_{n_w}^n v'(x(t)) \frac{dx(t)}{d\lambda_E} \frac{d\lambda_E}{db} dt.$$
$$(44)$$

Substituting $v'(x_w) = 1/[n_w(1 - \tau(n_w)]$ to eliminate $v'(x_w)$, we find that the first term at the right-hand side of this equation is positive. Also the second term at the right-hand side is positive if $d\lambda_E/db < 0$ (since (12) implies that $\frac{dx(t)}{d\lambda_E} < 0$). Hence, $d\lambda_E/db < 0$ together with $\frac{d\tau(n)}{dn} < 0$ around $n = n_w$ implies that all agents gain from a higher welfare benefit.

(44) implies that a value n^* close to n_w exists such that $\frac{du(n)}{db} > 0$ for all $n \in \langle n_w, n^* \rangle$. Since $\frac{d\lambda_E}{db} > 0$ implies that $\frac{dx(n)}{db} < 0$ for all $n > n_w$, $\frac{du(n)}{db} > 0$

implies that $\frac{dz(n)}{db} < 0$ for all $n \in \langle n_w, n^* \rangle$. Finally, rewriting (42) for the effect of b on $T(n)$, we obtain

$$\frac{dT(n)}{db} = \underbrace{-n\frac{du(n)}{db}}_{<0} + \underbrace{[nv'(x(n)) - 1]}_{= \frac{\tau(n)}{1-\tau(n)} \geq 0}\underbrace{\frac{dx(n)}{db}}_{<0} < 0$$

for all $n \in \langle n_w, n^* \rangle$, where we have used that $\tau(n_w) > 0$ and hence $\tau(n) > 0$ for n close to n_w. Q.E.D.

Proof of proposition 6

Consider the case in which $b + T(n_w) < 0$ and $\tau(n_w) < 0$. The first-order condition for n_w (see (24) with $u(n_w) = v(b) + \gamma$) then implies that $\eta_w < 0$.

Employing (42), we find for types n with $\tau(n) < 0$ that

$$\frac{dT(n)}{dE} = \underbrace{-n\frac{du(n)}{dE}}_{>0} + \underbrace{[nv'(x(n)) - 1]}_{= \frac{\tau(n)}{1-\tau(n)} < 0}\underbrace{\frac{dx(n)}{dE}}_{<0} > 0,$$

where the effects of E on $u(n)$ and $x(n)$ are derived in lemma 4. Q.E.D.

Figure 1: equilibrium in n_w, λ_E space is determined by the intersection of the Labor Supply (LS) and Government Budget Constraint (GBC) curves. The LS curve should be nondecreasing and hence must be ironed out for $n_w < \underline{n}$. Government tax revenues are maximized if $n_w = \tilde{n}_w$.