WORKING PAPER 2025:19

Redistribution and labor market inclusion

Thomas Aronsson
Spencer Bastani
Khayyam Tayibov

IFAU
INSTITUTE FOR EVALUATION OF LABOUR MARKET AND EDUCATION POLICY

The Institute for Evaluation of Labour Market and Education Policy (IFAU) is a research institute under the Swedish Ministry of Employment, situated in Uppsala.

IFAU's objective is to promote, support and carry out scientific evaluations. The assignment includes: the effects of labour market and educational policies, studies of the functioning of the labour market and the labour market effects of social insurance policies. IFAU shall also disseminate its results so that they become accessible to different interested parties in Sweden and abroad.

Papers published in the Working Paper Series should, according to the IFAU policy, have been discussed at seminars held at IFAU and at least one other academic forum, and have been read by one external and one internal referee. They need not, however, have undergone the standard scrutiny for publication in a scientific journal. The purpose of the Working Paper Series is to provide a factual basis for public policy and the public policy discussion.

More information about IFAU and the institute's publications can be found on the website www.ifau.se

ISSN 1651-1166

Redistribution and labor market inclusion*

Thomas Aronsson† Spencer Bastani‡ Khayyam Tayibov§

October 14, 2025

Abstract

This paper incorporates labor market inactivity and long-term unemployment into the framework of optimal redistributive taxation. We examine how a combination of education policy, public employment programs, unemployment benefits, and optimal income taxation can effectively address both redistributive goals and the persistent challenges of long-term unemployment. Our analysis shows that the second-best optimal policy typically implies overprovision of education compared to a policy rule that reflects only direct marginal benefits and costs. At the same time, public employment programs and unemployment benefits tend to be underprovided. Through numerical simulations, we illustrate how this policy mix adapts to varying preferences for redistribution, productivity disparities, and the proportion of individuals at risk of long-term unemployment.

Keywords: long-term unemployment, education, optimal income taxation, public sector employment

JEL classification: H21, J24, J45, I21

*We are grateful for comments from Floris Zoutman, Etienne Lehmann, Katarina Nordblom, Håkan Selin, Tobias König, Thomas Giebe, and Åsa Hansson. The authors would like to thank The Marianne and Marcus Wallenberg Foundation (MMW 2015.0037) and the Swedish Research Council (2020-02208) for generous research grants.

†Addresses: (i) Department of Economics, Umeå School of Business, Economics and Statistics, Umeå University, SE-901 87 Umeå, Sweden; and (ii) Department of Economics, University of Graz, A-8010 Graz, Austria. E-mail: thomas.aronsson@umu.se

‡Institute for Evaluation of Labour Market and Education Policy (IFAU) and Department of Economics, Uppsala University; Research Institute for Industrial Economics (IFN), Stockholm, Sweden; Uppsala Center for Fiscal Studies (UCFS), Uppsala Center for Labor Studies (UCLS), CESIfo, Germany. E-mail: spencer.bastani@ifau.uu.se.

§Department of Economics and Statistics, School of Business and Economics, Linnaeus University, Växjö, Sweden. E-mail: khayyam.tayibov@lnu.se.

1 Introduction

Long-term unemployment is a major challenge in today's labor markets, which are characterized by rapid changes in job structures, increasing specialization, and rising skill requirements. This problem is exacerbated by transnational migration patterns and the limited labor market integration of certain immigrant groups. In addition, economic shocks (such as the global COVID-19 pandemic) can exacerbate the problem of long-term unemployment by hindering the reintegration of displaced workers. A key question arises: How should tax systems and public spending programs respond to these challenges? This paper aims to answer this question by analyzing the policy implications of long-term unemployment, focusing on the design of optimal income taxation, education policies, and public employment training programs.

To this end, we develop and examine a discrete, intertemporal adaptation of the Mirrlesian Mirrlees (1971) optimal income tax model that includes a nonlinear income tax, transfers to individuals engaged in education, a public employment program with predetermined wage rates and effort requirements, and unemployment benefits.[1] We combine theoretical analysis, focusing on the characterization of policy rules for marginal taxation and the aforementioned public expenditure programs, with numerical simulations illustrating the optimal policy responses under a range of economic scenarios.

The basic elements of our model are as follows: Individuals optimize over two distinct periods, with employment opportunities available in both periods to those who are sufficiently productive. We classify individuals into three skill types, denoted by skill types 1, 2, and 0. Skill types 1 and 2 represent conventional low-skilled and high-skilled individuals, respectively, whose pre-tax wage rates are determined by their intrinsic productivity throughout their working lives. We introduce a novel skill type 0 characterized by market productivity below its innate potential, which renders this type uncompetitive in the regular labor market.[2] Our primary focus is on type 0 agents and on policies to combat long-term unemployment, where the qualitative insights do not depend on the number of latent skill types.

In addition to differing in their latent ability, individuals of skill type 0 can also differ in their effort costs of education and can choose between unemployment in the first period, in which case they will remain unemployed throughout their working lives, a public employment program, or education. Education and the public employment program allow individuals of type 0 to raise their market productivity sufficiently to secure regular employment in the second period. While education allows them to reach

[1]Empirical research supports the idea that public employment programs can improve the employment prospects of groups with weak labor market attachment (see, for example, Mörk et al. (2021), for a recent evaluation of a temporary public employment program in Sweden).

[2]Our model builds on and extends the discrete models of optimal nonlinear income taxation originally developed by Stern (1982) and Stiglitz (1982).

their true earning potential, a public employment program raises their productivity just enough to enter the regular labor market in the second period. Crucially, our model emphasizes that inaction by individuals with skill type 0 in the first period leads to long-term unemployment.[3] This is costly both to the individual and to society. The social objective is represented as a weighted combination of individual utilities, with weights reflecting preferences for redistribution, and a societal inclination to alleviate long-term unemployment.

The results show that the policy rules for marginal income taxation are broadly similar to those in models without unemployment. However, an important difference is that these policy rules take into account the endogeneity of the skill distribution within the labor force, which is influenced by government policies aimed at labor market integration. The marginal income tax policy does not directly affect the discrete activity choices of individuals of skill type 0, since their decisions are based on utility comparisons across alternatives and thus on total tax contributions rather than on marginal incentives.

The policy rules for the wage rate and effort requirements in the public employment program, education transfers, and unemployment benefits all depend on how each of these instruments affects the activity choices of type 0 individuals and thus the present value of the net tax revenue available for redistribution. For example, if in practice the present value of net tax revenue is higher when individuals with skill type 0 choose education over the public employment program, then, all else equal, tax revenue considerations will require (i) a lower wage in public employment and (ii) higher education transfers compared to scenarios in which these two activities yield equivalent present values of net tax revenue. Tax revenue concerns also impose constraints on unemployment benefits, reflecting society's aversion to prolonged unemployment.

Our numerical analysis explores the dynamics of resource allocation and public policy by examining the responses to variations in the government's propensity to redistribute, changes in the dispersion of the productivity distribution, and variations in the share of skill type 0 agents in the economy. This analysis shows that an increase in the government's propensity to redistribute and a mean-preserving increase in the dispersion of the productivity distribution lead to more generous unemployment benefits and increased long-term unemployment, especially among the latent low-skilled. An increased share of agents with skill type 0 leads to more public employment among low-skilled, accompanied by lower educational attainment and a slight reduction in unemployment. The activity choices of the high-skilled show less sensitivity in this context. Moreover, each of these changes tends to increase the cost of redistribution in equilibrium, leading to higher marginal taxation for the low-skilled and higher average taxation for the

[3] Kieselbach (2003), in a study of long-term youth unemployment in the EU, identifies low skills and labor market inactivity among unemployed youth as factors that increase their risk of social exclusion. Our model captures these characteristics, as low market productivity combined with inactivity early in life leads to exclusion from the labor market.

high-skilled.

The present study contributes to the literature in several ways. First, it represents the first attempt to integrate labor market inactivity and its possible consequences in terms of labor market exclusion into the theory of optimal redistributive taxation and public spending. While previous studies have addressed related issues such as poverty alleviation (e.g., Pirttilä and Tuomala (2004); Kanbur et al. (2018)) and unemployment (e.g., Marceau and Boadway (1994); Lehmann et al. (2011); Aronsson and Sjögren (2004); Hungerbühler et al. (2006); Aronsson and Johansson-Stenman (2021); Aronsson and Micheletto (2021)), none has examined the consequences of labor market inactivity and its policy implications in an optimal tax framework.

Second, our study contributes to research on optimal income taxation and occupational choice, building on the seminal work of Diamond (1980), Boadway et al. (1991), and Saez (2002). Recent work in this area has examined the equilibrium effects of sectoral reallocation of labor on wages, as in Rothschild and Scheuer (2013), and the externalities arising from occupational choice, as in Rothschild and Scheuer (2016) and Lockwood et al. (2016).[4]

Third, our research explores the joint design of income tax and education policies to promote human capital development and learning-by-doing. This is consistent with the work of Bovenberg and Jacobs (2005), Maldonado (2008), Findeisen and Sachs (2016), Colas et al. (2021), Stantcheva (2017), and Da Costa and Santos (2018). In contrast to these studies, our analysis emphasizes the validation of skills rather than the creation of new human capital.

Fourth, weak attachment to the labor market and the negative consequences thereof for individuals' abilities to control their own economic resources are key aspects of the broader concept of social exclusion (e.g., Atkinson (1998); Sen (2000); Bradley et al. (2003); Pohlan (2019)). Our study contributes to research on social exclusion by examining the policy implications of labor market inactivity and how an optimal redistribution policy should be modified when certain groups of individuals are at risk of labor market exclusion.

The structure of the paper is as follows: Section 2 presents the model and outlines the private decision problems, while section 3 details the public decision problem. Section 4 characterizes the optimal policy rules for marginal tax policy, education policy, public employment program, and unemployment benefits. The numerical model is outlined in Section 5, and numerical results and sensitivity analyses are discussed in Section 6. Finally, concluding remarks are given in section 7. The proofs are given in Appendix A, and additional mathematical and numerical results are available in the Online Appendix.

[4] Our paper is also related to the recent strand of the literature focusing on optimal tax responses to new technologies and the optimal taxation of "robots," see, for example, Guerreiro et al. (2022), Costinot and Werning (2023), and Thuemmel (2023).

2 Model

Types of agents There are three types of individuals in the economy, labeled 0, 1, and 2, respectively. Each individual optimizes over two periods. Type 1 and 2 individuals are employed in both periods. Our main focus is on type 0 individuals who are not productive enough to obtain regular employment. One example is young migrants who have not yet acquired the necessary skills (such as language skills) for regular employment. Another is people who did not do well enough in school or did not finish school (for a variety of reasons, such as family background or peer group influence), which makes them less employable. As explained in more detail below, in the first period, type 0 agents choose between unemployment (U), education (E), and a public employment program (P). These activity choices, in turn, determine their market wage rates and employment opportunities in period 2.

Market productivity Individuals of types 1 and 2 (sometimes called "conventional" agents) are characterized by innate productivities θ^1 and θ^2, where $\theta^2 > \theta^1$, and have market productivities of $w^1 = \theta^1$ and $w^2 = \theta^2$. [5] In contrast, individuals of type 0 can be either of innate productivity θ^1 or θ^2, but their market productivities are not equal to their innate productivities. Instead, regardless of their innate productivity, type 0 individuals have a marginal market productivity of $\alpha\theta^1$ in the first period, where $0 < \alpha < 1$.[6]

Minimum wage There is a minimum wage in the regular labor market, w^1, defined as the market wage of type 1. Since $\alpha\theta^1 < w^1$ by assumption, type 0 individuals cannot find regular employment in period 1. On the other hand, if $\alpha\theta^1$ exceeds what we can think of as a "technological minimum wage" (in the sense that there are productive tasks given the technology of the economy) equal to $\underline{w} < w^1$, then a type 0 individual can participate in a public employment program. We assume that this condition is satisfied, which ensures that public employment is available to all individuals of type 0. The discrepancy between w^1 and $\underline{w}$ can be due to minimum wage legislation or unionization of the labor market, and is a key feature of modern welfare states facing the problem of weak labor market attachment among certain groups of the population. To simplify the

[5] We assume that firms can observe the productivity of workers, thus abstracting from the screening/signaling possibilities available to firms/workers analyzed by, for example, Stantcheva (2014), Bastani et al. (2015), Craig (2023), Sztutman (2024) and Bastani et al. (2025). We also abstract from labor market discrimination (see, e.g., Blumkin et al. (2007)).

[6] An alternative would be to assume that the market productivity in the first period is given by $\alpha\theta^i$, $i = 1, 2$. This would open up the possibility that type 0 agents with latent productivity θ^2 could get a job with a market wage of $\alpha\theta^2$ in the first period, provided $\alpha\theta^2 \geqslant w^1$. Such an extension complicates the analysis considerably without having a major impact on the qualitative results.

analysis, w^1 is assumed to be exogenous in the following.[7]

Effort costs In addition to differing in their innate productivity θ, type 0 individuals also differ in their effort cost of training ξ, which is assumed to be continuously distributed with positive support. The parameter ξ will play a key role in determining whether innately high productivity individuals find it optimal to pursue education and thereby realize their true innate productivity, or to settle for public employment and lower future career prospects.

Tax instruments Since we are interested in analyzing policy problems faced by modern welfare states, it is important to capture the potential flexibility of the tax and transfer system. We assume that the government uses a labor income tax, given by general nonlinear tax functions $T_j(\cdot)$, $j = 1, 2$, where tax payments can be either positive or negative (the latter representing a transfer payment). The functions T_1 and T_2 represent the labor income taxes paid in periods 1 and 2, respectively. Since the tax functions T_1 and T_2 are not necessarily the same, we do not rule out age-dependent labor income taxation.[8]

Conventional agents Individuals of types 1 and 2 are employed in the regular labor market in both periods and solve the following problem:

$$W(\theta^i) = \max_{h_1,h_2} \left\{ u(\theta^i h_1 - T_1(\theta^i h_1)) - v(h_1) + \beta[u(\theta^i h_2 - T_2(\theta^i h_2)) - v(h_2)] \right\}, \quad (1)$$

where h_1 and h_2 denote the labor supplied in periods 1 and 2, respectively, and β denotes the utility discount factor. The function $u(\cdot)$ is the utility derived from consumption and is assumed to be twice differentiable, increasing, and strictly concave, which $v(\cdot)$ is a twice continuously differentiable, increasing and strictly convex function measuring the disutility of effort.

2.1 Activity choices

We now describe the activity choices available to type 0 agents in the first period.

[7] One way of motivating this assumption is that the minimum wage is outside the control of the national government, which is a relevant scenario in Europe in light of the recently proposed EU minimum wage directive. Another is that wage setting is centralized through nationwide trade unions.

[8] In principle, the optimal income tax could also be made to depend on the agent's income history. However, this makes little sense in our model, since individuals with market productivity θ^i $(i = 1, 2)$ in the second period will behave identically and be equally well off for the rest of their lives, regardless of whether they were a conventional type i or a type 0 in the first period.

Public employment If a type 0 individual is employed in the public sector in the first period, her marginal productivity is given by $\alpha\theta^1$ and she receives a wage, w_P, in return for a fixed labor supply of h_P units of time. We will later treat w_P and h_P as control variables of the government. The wage in public employment satisfies $w^1 \geqslant w_P > \alpha\theta^1$. Thus, a key feature of public employment is that workers are paid above their marginal product (a private firm would not offer such a contract as it would yield negative profits), but not above the minimum wage in regular employment. An alternative to public employment would be to introduce wage subsidies to private firms hiring type 0 workers. However, such subsidies would be strictly dominated by public employment for at least two reasons. First, firms have strong incentives to claim wage subsidies for ordinary workers that they would be willing to hire also in the absence of a subsidy. Second, unlike private firms, the government can absorb the potential risk of hiring a worker with below-market productivity at no cost.

In the second period, after working in public employment, the marginal market productivity increases from $\alpha\theta^1$ to θ^1, and the agent works in a regular job with a labor supply of h_2.[9] Formally, the indirect utility of an agent of type i with training cost ξ choosing public employment is:

$$V_P^i(\xi) = \max_{h_2} \left\{ u(w_P h_P) - \xi v(h_P) + \beta[u(\theta^1 h_2 - T_2(\theta^1 h_2)) - v(h_2)] \right\}. \tag{2}$$

All individuals entering public employment in the first period (regardless of their innate productivity) will earn the market wage θ^1 in the second period. This means that individuals of type 0 with innate productivity θ^2 do not realize their maximum potential productivity. This is quite natural, since the goal of public employment is to learn the basic skills required for regular employment (e.g., gaining work experience, establishing a social network, and acquiring language skills); they do not acquire the more advanced skills required for high-skill jobs.

Education The purpose of education is to enable agents to transform their innate skills into marketable skills, thereby realizing their true innate productivity. One interpretation of education is that it serves to *increase* an individual's skill level; another is that it serves to *validate* those skills, as in enabling an immigrant with a foreign medical degree to practice medicine in the host country.

Formally, the indirect utility of an individual of type 0 with innate productivity θ^i enrolling in education in the first period is given by

$$V_E^i(\xi) = \max_{h_2} \left\{ u(c_E) - \xi v(e^i) + \beta \left[u(\theta^i h_2 - T_2(\theta^i h_2)) - v(h_2) \right] \right\}, \tag{3}$$

[9]Thus, public employment has a learning-by-doing component, see, for example, Stantcheva (2017) and Makris and Pavan (2021).

where e^i is the effort required by type 0 with innate productivity i to realize his/her innate productivity through education.

Note that while education is available to both types of innate productivity, it serves, nevertheless, a primary purpose for those with innate productivity θ^2. This is because the innate high skilled are only able to realize their (high) productivity through education. As a result, it is typically optimal for the government to incentivize type 0 agents with latent productivity θ^2 to choose education, unless θ^2 is small relative to θ^1 so that the productivity gains of education are modest.

For agents with innate productivity θ^1, their period-2 productivity is the same regardless of whether they choose public employment or education. While these agents may still strictly prefer education due to its lower effort requirement, from the government's perspective it is generally optimal to structure taxes and transfers so that most θ^1-type agents choose public employment. This approach ensures that these agents contribute to output immediately, rather than producing nothing during training.

Unemployment Finally, if a type 0 individual chooses to become unemployed in the first period, this choice will lead to long-term unemployment, i.e., unemployment also in the second period. In the case of unemployment, the agent's indirect utility is given by

$$V_U = u(b_1) + \beta u(b_2), \tag{4}$$

where b_1 and b_2 is a profile of unemployment benefits decided by the government. It is important to keep in mind that although individuals can be viewed as "choosing" unemployment, it is not really a conventional choice, since the agents who choose unemployment are those with the highest effort costs, who are least able to escape long-term unemployment and achieve economic advancement.

Although unemployment may allow individuals to spend more time on leisure activities, some previous studies show that unemployment leads to a loss of well-being over and above the income loss it causes (e.g., Clark and Oswald 1994; Winkelmann and Winkelmann 1998). In principle, we could capture this utility loss by adding an additional term to the equation (4). Such a cost would not change the analysis, since the choice between unemployment and training depends on their *relative* attractiveness, which is influenced by the parameter ξ. Instead, the additional social cost of unemployment is captured by an extra term in the social welfare function, which reflects society's aversion to long-term unemployment (see subsection 3.1).

3 The policy problem

The government (or social planner) collects revenue and redistributes through a nonlinear income tax as well as provides a public employment program, education transfers, and unemployment benefits. The government can observe income at the individual level, while latent skills and effort costs are private information. Thus, the government can neither verify the educational effort nor the labor supply. These are standard assumptions in Mirrleesian models of optimal income taxation. We assume that the government is first mover vis-à-vis the private sector and is able to commit to the optimal tax and expenditure policy, which is decided before individuals make their choices.[10]

Access to general income taxes means that the government can implement any desired combination of consumption and labor supply in each period for both latent skill types, subject to incentive compatibility constraints and the public budget constraint. Therefore, according to the Taxation Principle (Hammond 1979), we can equivalently let the government choose these allocations directly, using incentive compatibility mechanisms such that each individual (weakly) prefers the allocation intended for her type over the allocations intended for other types. The policy rules for marginal taxation can then be derived by combining social and private first-order conditions. The other policy instruments, i.e., education policy, public employment (wage and hours requirements), and unemployment benefits, are chosen directly by the government.

Let (y_t^i, c_t^i) denote the pre-tax income/consumption bundles in period $t = 1, 2$ offered to individuals of type $i = 1, 2$. For individuals of types 1 and 2, who realize their innate productivity immediately and are, therefore, in regular employment throughout their working lives, utility is given by

$$W^i = u(c_1^i) - v(y_1^i/\theta^i) + \beta\left(u(c_2^i) - v(y_2^i/\theta^i)\right), \quad i = 1, 2, \tag{5}$$

if they choose the bundles intended for them by the government.

Similarly, the utilities of type 0 individuals who choose the bundles intended for them are given as follows depending on their activity choice:

$$V_P^i(\xi) = u(w_P h_P) - \xi v(h_P) + \beta\left(u(c_2^1) - v(y_2^1/\theta^1)\right), \quad i = 1, 2 \tag{6}$$

$$V_E^i(\xi) = u(c_E) - \xi v(e^i) + \beta\left(u(c_2^i) - v(y_2^i/\theta^i)\right), \tag{7}$$

$$V_U^i(\xi) = u(c_1^U) + \beta u(c_2^U), \quad i = 1, 2 \tag{8}$$

where $\{c_1^U, c_2^U\}$ denotes the consumption profile of the unemployed. Combining the

[10] The assumption of full commitment greatly simplifies the analysis in an already rather complex model, although we are aware of the potential time inconsistency problem that this commitment raises. Aronsson and Sjögren (2016) and Brett and Weymark (2019) analyze time-consistent optimal taxation in different contexts under asymmetric information.

above equations, the utility of an agent of type (θ^i, ξ) can be summarized as

$$V^i(\xi) = \max\{V^i_P(\xi), V^i_E(\xi), V^i_U(\xi)\}, \quad i = 1, 2. \tag{9}$$

We assume that the government wants to redistribute from high-productivity to low-productivity individuals. Thus, the government must prevent high-productivity individuals from mimicking low-productivity individuals in order to benefit from this redistribution. In our framework, there are two such ways. First, to ensure that type 0 individuals who realize the productivity level θ^2 in the second period (i.e., those with innate productivity θ^2 who choose education) weakly prefer the allocation intended for them to the allocation intended for the low-productivity type, we must impose the following self-selection constraint:

$$u(c_E) - \xi v(e^2) + \beta[u(c^2_2) - v(y^2_2/\theta^2)] \geqslant u(c_E) - \xi v(e^2) + \beta[u(c^1_2) - v(y^1_2/\theta^2)]. \tag{10}$$

Since the first two terms on either side of the inequality are identical, the constraint can be simplified to

$$u(c^2_2) - v(y^2_2/\theta^2) \geqslant u(c^1_2) - v(y^1_2/\theta^2). \tag{11}$$

Second, we need a self-selection constraint for conventional type 2 individuals, i.e., those with realized productivity θ^2 in the in the first and second periods, so that they (weakly) prefer the allocation intended for them to the allocation intended for a conventional type 1:

$$u(c^2_1) - v(y^2_1/\theta^2) + \beta[u(c^2_2) - v(y^2_2/\theta^2)] \geqslant u(c^1_1) - v(y^1_1/\theta^2) + \beta[u(c^1_2) - v(y^1_2/\theta^2)]. \tag{12}$$

If (11) is binding, then constraint (12) is simplified to

$$u(c^2_1) - v(y^2_1/\theta^2) \geqslant u(c^1_1) - v(y^1_1/\theta^2). \tag{13}$$

In the rest of the theory section, we will assume for simplicity that (11) is binding. This will always be the case if there is sufficient redistribution in the second period. We therefore include (11) as an equality and (13) in the social decision problem.[11]

In addition to (11) and (13), we want to prevent *conventional* type 1 and type 2 workers from mimicking type 0 in order to benefit from the education and public employment policies. Admittedly, it is not entirely clear how to formulate these constraints, since the conventional types already have the skills needed to enter the regular labor

[11] In the numerical simulations, we impose (11) as a weak inequality and (12) in its full form, since we cannot guarantee that (11) is binding under all possible parameter constellations. The relatively simple structure of the incentive constraints is facilitated by the fact that type 0 workers who realize productivity θ^2 in the second period behave identically to type 2 workers in the second period.

market. To this end, we make three observations based on the assumptions made above. First, conventional type 1 and type 2 agents entering public employment must supply h_P hours of labor in the first period and will earn the same compensation as type 0 workers in public employment ($w_P h_P$). Second, type 1 and type 2 individuals who enter education to mimic type 0 are required to make a minimal effort, $\underline{e}$, to be eligible for the education transfer c_E.[12] Third, the benefit that a conventional type 1 or 2 agent derives from mimicking type 0 is entirely confined to the first period, since it is assumed that these mimickers obtain employment according to their actual (already realized) productivity in the second period.[13]

Given the above discussion, to prevent a conventional type 1 worker from mimicking type 0 in education or public employment, we require that he/she obtains a weakly higher utility in period 1 when working in the regular labor market by imposing the following constraints:

$$u(c_1^1) - v(y_1^1/\theta^1) \geqslant u(w_P h_P) - v(h_P) \tag{14}$$

$$u(c_1^1) - v(y_1^1/\theta^1) \geqslant u(c_E) - v(\underline{e}). \tag{15}$$

Note that (14) and (15) are the only conditions we need to impose, since they also prevent conventional type-2 agents from mimicking type-0. The reason is that $u(c_1^2) - v(y_1^2/\theta^2) > u(c_1^1) - v(y_1^1/\theta^1)$ in a second-best optimum.[14]

3.1 Social decision-problem

Let $f_{\xi|\theta}(\xi \mid \theta)$ denote the conditional probability distribution of ξ given θ. The joint probability distribution then becomes $f_{\xi,\theta}(\xi, \theta) = f_{\xi|\theta^1}(\xi \mid \theta = \theta^1)\pi^1 + f_{\xi|\theta^2}(\xi \mid \theta = \theta^2)\pi^2$, where π^1 and π^2 are the fractions of type 0 agents with innate productivity θ^1 and θ^2, respectively. To shorten the notation, define $f_{\xi|\theta^1}(\xi \mid \theta = \theta^1) = f^1(\xi)$ and $f_{\xi|\theta^2}(\xi \mid \theta = \theta^2) = f^2(\xi)$, with corresponding CDFs F^1 and F^2, and let γ^0, γ^1, and γ^2 be the fractions of types 0, 1, and 2, respectively, in the population, where $\sum_i \gamma^i = 1$.

We assume that the social objective function consists of two parts. The first part is welfarist and is given by a generalized utilitarian objective. The second part is non-

[12] The parameter $\underline{e}$ should be viewed as the minimum effort required to be eligible for transfers from the government, assuming that some monitoring is possible (e.g., an attendance requirement).

[13] Given that educational effort could in principle serve as a signal of innate ability, we assume that the minimal effort level $\underline{e}$ is sufficiently large to imply that firms are not deterred from compensating conventional type 1 and 2 workers according to their actual (already realized) productivity in the second period.

[14] Note that self-selection constraints as we have formulated them do not prevent conventional type 1 and type 2 workers from mimicking the unemployed. Since such a scenario seems less likely to us, and our framework is already quite complex, we abstract from this possibility when formalizing the model. We will, nevertheless, return the implications of such incentives when the optimal unemployment benefits are characterized in subsection 4.4.

welfarist and depends directly on the number of long-term unemployed, aiming to capture the desire to reduce the overall social cost of long-term unemployment.[15] Formally, the social objective is given by

$$\phi \int_0^\infty \Big[\sum_{i=1,2} \pi^i f^i(\xi) V^i(\xi) \Big] d\xi + \sum_{i=1,2} \gamma^i W^i + \sum_{i=1,2} H^i \left(\gamma^0 \pi^i \int_{\Omega_U^i} f^i(\xi) d\xi \right), \tag{16}$$

where V and $W^i, i = 1,2$ are defined in (9) and (5), respectively, and ϕ denotes the welfare weight the government attaches to type 0 agents. The H^i functions can be interpreted in terms of social aversion to long-term unemployment, and this aversion is allowed to vary depending on whether the long-term unemployed are latently high-skilled or low-skilled.[16] We assume that H^i is strictly decreasing and concave in the number of long-term unemployed in society, $\gamma^0 \pi^i \int_{\Omega_U^i} f^i(\xi) d\xi$, for $i = 1,2$. The welfarist part of (16), i.e., the sum of the first two terms, is consistent with a utilitarian social welfare function in the special case where $\phi = \gamma^0$, while $\phi > \gamma^0$ represents a case where more weight is given to the welfare of type 0 from an ex ante perspective.

The government's problem is to choose the pre-tax income and consumption variables $\{y_1^i, y_2^i, c_1^i, c_2^i\}_{i=1,2}$, the education transfer c_E, the time path of unemployment benefits $\{c_1^U, c_2^U\}$, and the wage rate and hours in public employment, $\{w_P, h_P\}$, to maximize social objective (16) subject to the incentive compatibility constraints (11) and (13)–(15) and the following public resource constraint (assuming, for simplicity, that the interest rate is zero):

$$\begin{aligned}
&\sum_{i=1,2} \gamma^i \Big[(y_1^i - c_1^i) + (y_2^i - c_2^i) \Big] + \gamma^0 \sum_{i=1,2} \pi^i \left(\int_{\Omega_E^i} f^i(\xi) d\xi \right) \left(-c_E + (y_2^i - c_2^i) \right) \\
&+ \gamma^0 \left(\sum_{i=1,2} \pi^i \int_{\Omega_P^i} f^i(\xi) d\xi \right) \left([\alpha\theta^1 - w_P] h_P + (y_2^1 - c_2^1) \right) - \\
&\gamma^0 \left(\sum_{i=1,2} \pi^i \int_{\Omega_U^i} f^i(\xi) d\xi \right) \left(c_1^U + c_2^U \right) \geqslant 0.
\end{aligned} \tag{17}$$

In (17), Ω_E^i, $i = 1,2$, is the set of type 0 individuals with innate productivity θ^i who enter education (and realize their true innate productivity). Formally, $\xi \in \Omega_E^i$ implies $V^i(\xi) = V_E^i(\xi)$. Similarly, the sets Ω_P^1 and Ω_P^2 are the sets of type 0 individuals of innate

[15] More general social welfare weights along the lines of Saez and Stantcheva (2016), which depend on both income and non-income characteristics, could easily be introduced. By focusing on a social welfare function with non-welfaristic elements, our analysis is related to recent work on fairness and poverty reduction by Henry de Frahan and Maniquet (2021) and Maniquet and Neumann (2021).

[16] One reason for social aversion against long-term unemployment would be the broader negative consequences this may have in terms of social exclusion and/or increased criminality. Note also that with a slight reformulation of the model, an alternative interpretation of the H^i functions would be that they represent social preferences at the individual level, i.e., that each individual attaches disutility to the severity of the problem of long-term unemployment.

productivity θ^1 and θ^2, respectively, who enter public employment in the first period. Therefore, $\xi \in \Omega_P^i$ implies $V^i(\xi) = V_P^i(\xi)$. In the last row, Ω_U^i is the set of type 0 individuals with innate productivity θ^i who choose unemployment in period 1, leading to long-term unemployment, i.e., exclusion from the labor market; thus, $\xi \in \Omega_U^i$ implies $V^i(\xi) = V_U^i(\xi)$ (the notation can be simplified if the relevant sets are intervals).

The resource constraint reflects that gross income (or output) is used for private consumption. Each term on the left-hand side can be interpreted as the net tax revenue collected from different groups of individuals over the two periods. The first term reflects the net tax revenue collected from individuals of types 1 and 2, respectively, while the remaining terms refer to individuals of type 0, where the public revenue collected depends on the choices made by these agents in the first period. Specifically, the second term measures the net tax revenue raised from type 0 individuals who enter education, while the third term correspondingly measures the net tax revenue raised by type 0 individuals who choose public employment. Finally, the fourth term reflects the fiscal cost associated with the consumption profile of the unemployed.

3.2 Activity choices and their ordering

The characterization of activity choices among type 0 individuals is based on a comparison of the indirect utility levels in equations (6)–(8). In our framework, depending on ξ (from low to high), individuals choose activity in the first period in the following order: public employment, education, and unemployment.

While it is natural to assume that those with the highest effort cost in training will choose unemployment, the ordering of public employment and education requires some discussion. Specifically, we assume that $w_P h_P > c_E$ and $v(h_P) > v(e^i)$, $i = 1, 2$, at the second best optimum. The first inequality is based on the observation that education is an investment that pays off in the future. The second can be interpreted to reflect the idea that education is associated with a consumption value, such that the disutility of educational effort is effectively smaller than the disutility of routine work.

Conditional on other characteristics, we can now present intervals for ξ associated with different activities, as well as critical values where an individual is indifferent between two distinct activities. For type 0 individuals with innate productivity θ^1, the choice between public employment and education is guided by the following utility difference:

$$V_P^1(\xi) - V_E^1(\xi) = u(w_P h_P) - u(c_E) - \xi(v(h_P) - v(e^1)).$$

We can then derive the effort cost for an individual who is indifferent between public employment and education by setting $V_P^1(\xi) = V_E^1(\xi)$ and then solving for ξ:

$$\xi_{P-E}^1 = \frac{u(w_P h_P) - u(c_E)}{v(h_P) - v(e^1)} > 0. \tag{18}$$

Therefore, whenever $\xi < \xi_{P-E}$ we have $V_P > V_E$, implying that the individual prefers public employment to education, while $\xi > \xi_{P-E}$ correspondingly implies $V_E > V_P$, such that the individual instead prefers education. We can derive a similar threshold for individuals with innate productivity θ^2. By setting $V_P^2(\xi) = V_E^2(\xi)$, we obtain:

$$\xi_{P-E}^2 = \frac{[u(w_P h_P) - u(c_E)] + \beta\left([u(c_2^1) - u(c_2^2)] - [v(y_2^1/\theta^1) - v(y_2^2/\theta^2)]\right)}{v(h_P) - v(e^2)}. \tag{19}$$

Contrary to (18), we can see that variables referring to the second period enter the right-hand side of (19). The reason for this discrepancy is that the choice between public employment and education affects the utility in period 2 (not just the utility in period 1) for agents with innate productivity θ^2.

Finally, for the choice between education and unemployment, the utility difference is given by (for i=1,2):

$$V_E^i(\xi) - V_U^i(\xi) = [u(c_E) - \xi v(e^i)] - u(c_1^U) + \beta\left([u(c_2^i) - v(y_2^i/\theta^i)] - u(c_2^U)\right). \tag{20}$$

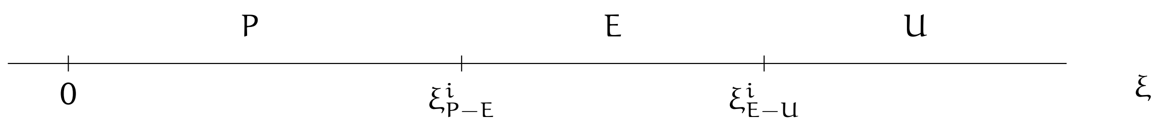
By setting $V_E^i(\xi) = V_U^i(\xi)$, the following critical (threshold) effort cost can then be derived:

$$\xi_{E-U}^i = \frac{1}{v(e^i)}\left([u(c_E) - u(c_1^U)] + \beta\left([u(c_2^i) - v(y_2^i/\theta^i)] - u(c_2^U)\right)\right). \tag{21}$$

Equation (21) means that a type 0 individual of innate productivity θ^i will prefer unemployment to education if $\xi > \xi_{E-U}^i$, and vice versa if $\xi < \xi_{E-U}^i$. We can see that the number of long-term unemployed with innate productivity θ^i depends positively on the educational effort and negatively on the income loss associated with unemployment. Note also that some unemployment always remains at the social optimum. This is because, with an unbounded effort cost distribution, there will be some individuals who suffer enormous disutility of effort. Thus, it is not optimal for the government to activate these individuals in the first period.

The order of activity choices, verified by our numerical simulations below, is illustrated in figure 1.

Figure 1: Occupational choice thresholds for type 0 individuals with innate productivity θ^i.



Some useful notation For later use, we introduce the following short notation:

$$N^i_E = |\Omega^i_E| = \gamma^0 \pi^i \left(F^i(\xi^i_{E-U}) - F^i(\xi^i_{P-E})\right) \tag{22}$$

$$N^i_P = |\Omega^i_P| = \gamma^0 \pi^i F^i(\xi^i_{P-E}) \tag{23}$$

$$N^i_U = |\Omega^i_U| = \gamma^0 \pi^i [1 - F^i(\xi^i_{E-U})] \tag{24}$$

Here, N^i_E, N^i_P, and N^i_U denote the number of type 0 individuals with innate productivity θ^i who enter education, PE, and unemployment, respectively, in the first period.

4 Optimal marginal tax and expenditure policy

We are now ready to characterize the optimal marginal tax policy and public expenditures. To this end, we use the social first-order conditions presented in Online Appendix A.3 and the private first-order conditions for the labor supply. For an individual with realized productivity θ^i, $i = 1, 2$, the private first-order condition for the labor supply in period $t = 1, 2$ can be written as follows:

$$\tau^i_{y,t}\theta^i \equiv \theta^i - \frac{v'(y^i_t/\theta^i)}{u'(c^i_t)}. \tag{25}$$

The left-hand side of equation (25) is the marginal labor income tax payment per unit of labor, measured as the marginal labor income tax rate, $\tau^i_{y,t}$, times the pre-tax wage rate per unit of labor, θ^i, and the right-hand side measures the discrepancy between the pre-tax wage rate and the marginal rate of substitution between leisure and private consumption. Thus, we can interpret equation (25) in terms of a tax wedge created by the labor income tax. Note that individuals of types 1 and 2 satisfy equation (25) in both the first and second periods, while individuals of type 0, who realize productivity θ^i (through an appropriate activity choice in the first period), satisfy equation (25) in the second period.

4.1 Marginal income taxes

Let μ_1 denote the Lagrange multiplier associated with self-selection constraint (13), which prevents a true type 2 from mimicking individuals of type 1, while μ_2 correspondingly denotes the Lagrange multiplier attached to self-selection constraint (11) preventing type 0 individuals of innate productivity θ^2 from mimicking type 1 in the second period. λ represents the Lagrange multiplier of the resource constraint (measuring the marginal cost of public funds in units of utility).

The marginal income tax rates implemented in the first and second period, respectively, are characterized in Proposition 1.

Proposition 1. *The marginal income tax rates in the first period are characterized by the following policy rules:*

$$\tau^1_{y,1} = \frac{\mu_1 u'(c^1_1)}{\lambda \gamma^1} \left(\frac{v'(y^1_1/\theta^1)}{u'(c^1_1)} \frac{1}{\theta^1} - \frac{v'(y^1_1/\theta^2)}{u'(c^1_1)} \frac{1}{\theta^2} \right) > 0, \tag{26}$$

$$\tau^2_{y,1} = 0, \tag{27}$$

while the marginal income tax rates implemented in the second period are characterized as follows:

$$\tau^1_{y,2} = \frac{\mu_2 u'(c^1_2)}{\lambda \left(\gamma^1 + \gamma^{0,1}\right)} \left(\frac{v'(y^1_2/\theta^1)}{u'(c^1_2)} \frac{1}{\theta^1} - \frac{v'(y^1_2/\theta^2)}{u'(c^1_2)} \frac{1}{\theta^2} \right) > 0, \tag{28}$$

$$\tau^2_{y,1} = 0, \tag{29}$$

where $\gamma^{0,1} = N^1_E + N^1_P + N^2_P$.

Proof See Appendix A.1. □

Proposition 1 reproduces an old result, albeit in a new context: the optimal marginal income tax rate implemented for the low-skilled is positive, while the marginal income tax facing the high-skilled is zero. We can also see from (26) that the policy rule governing the marginal tax treatment of the low-skilled takes a standard form in the first period.

In the second period, the number of low-skilled is endogenous, which means that the marginal tax treatment of this group is based on a slightly different policy rule. This can be seen from (28), in which $\gamma^1 + \gamma^{0,1}$ replaces γ^1, where $\gamma^{0,1}$ denotes the sum of type 0 individuals with (i) intrinsic productivity θ^1 enrolling in education in the first period, and (ii) individuals with intrinsic productivities θ^1 and θ^2 choosing PE in the first period. Thus, an important difference between (26) and (28) is the downward pressure on the marginal income tax in the second period following the enrollment in education and PE in the first period.

Finally, note that the incentives facing type 0 agents to engage in different activities in the first period are determined by comparisons of utility levels, which are, in turn, determined by total (not marginal) tax payments. Since the other policy instruments (c_E, w_P, h_P, c_1^U, and c_2^U) directly target these activity choices, there is no need to use marginal income taxation as an indirect instrument to influence them.

4.2 The optimal public employment program

To shorten the notation, let

$$\Gamma_E^i = -c_E + (y_2^i - c_2^i), i = 1, 2, \tag{30}$$

$$\Gamma_P = (\alpha\theta^1 - w_P)h_P + (y_2^1 - c_2^1) \tag{31}$$

denote the net tax revenue generated by an individual with innate productivity θ^i entering education and public employment, respectively, in the first period. Using μ_3 to denote the Lagrange multiplier attached to self-selection constraint (14), which serves to prevent a true type 1 from mimicking type 0 by entering public employment in the first period, the socially optimal wage and work effort in PE are characterized in Proposition 2.

Proposition 2. *Under optimal income taxation, the optimal wage w_P and work effort h_P in public employment satisfy, respectively, the following policy rules:*

$$\begin{aligned}&\left[\frac{\phi}{\gamma^0}u'(w_P h_P) - \lambda\right] h_P \left(N_P^1 + N_P^2\right) \\ &\qquad = \lambda\left(\sum_{i=1,2}\left(\frac{dN_E^i}{dw_P}\Gamma_E^i\right) - \frac{d\left(N_P^1 + N_P^2\right)}{dw_P}\Gamma_P\right) + \frac{\mu_3}{\gamma^0}u'(w_P h_P)h_P\end{aligned} \tag{32}$$

$$\begin{aligned}&\sum_{i=1,2} N_P^i\left[\frac{\phi}{\gamma^0}\left(w_P u'(w_P h_P) - v'(h_P)E_{f^i}(\xi \mid \xi \in \Omega_P^i)\right) - \lambda(w_P - \alpha\theta^1)\right] \\ &\qquad = \lambda\left(\sum_{i=1,2}\left(\frac{dN_E^i}{dh_P}\Gamma_E^i\right) - \frac{d\left(N_P^1 + N_P^2\right)}{dh_P}\Gamma_P\right) + \frac{\mu_3}{\gamma^0}\left(u'(w_P h_P)w_P - v'(h_P)\right)\end{aligned} \tag{33}$$

Proof See Appendix A.2. □

We will use the terms direct marginal benefit and direct marginal cost to denote those marginal benefits and costs that are neither driven by the activity choices in the first period nor by any binding self-selection constraint. The left-hand side of (32) and (33) thus measures the difference between the direct marginal benefit and the direct marginal cost of w_P and h_P, respectively, and reflects the extent to which each such instrument is overprovided or underprovides relative to a standard, first-best policy rule.[17]

Without any activity choices in the first period, i.e., if the number of individuals in each such group were fixed, the direct marginal benefit would be equal to the direct marginal cost at the social optimum. The discrepancy between them therefore reflects the fact that the wage rate and the hours requirement in public employment affect the activity choices of type 0 individuals (the terms proportional to λ on the right-hand side) as well as the incentives of type 1 individuals to mimic type 0 (the terms proportional to μ_3). Thus, the right-hand side of (32) and (33), respectively, is a consequence of the government wanting to influence these activity choices.

Let us start with the policy rule for the wage in (32). Note first that $\mu_3 \geqslant 0$, which means that the final term on the right-hand side helps to (weakly) reduce the wage in public employment below the level that equalizes the direct marginal benefit and the marginal cost. The intuition is, of course, that the government aims at preventing type 1 individuals from mimicking type 0 (which type 1 could otherwise do by entering public employment).

Turning to the terms proportional to λ in the second row, it is easy to see that $dN_P^i/dw_P > 0, i = 1, 2$. Therefore, given that $dN_E^i/dw_P = -dN_P^i/dw_P$, $i = 1, 2$, a sufficient condition for the second-best optimal wage to satisfy $\phi u'(w_P h_P)/\gamma^0 - \lambda > 0$ is

$$\Gamma_E^i = -c_E + (y_2^i - c_2^i) > [\alpha\theta^1 - w_P]h_P + (y_2^1 - c_2^1) = \Gamma_P, \quad i = 1, 2. \tag{34}$$

Thus, if condition (34) is satisfied, the right-hand side of (32) is unambiguously positive, in which case $\phi u'(w_P h_P)/\gamma^0 - \lambda > 0$. The intuition is straightforward: as long as the present value of net tax revenues is higher when type 0 chooses education rather than public employment, the government will set a lower wage in public employment. This reduces the number of persons in public employment and correspondingly increases the number of persons in education. In other words, public funds are costly and a lower wage in public employment leads to additional tax revenue, which, in turn, opens up for more redistribution.

To go further, suppose (realistically) that $\Gamma_E^2 > \Gamma_E^1$. Then (34) will automatically hold for $i = 2$ if it holds for $i = 1$. In this case, (34) is equivalent to

$$c_E < [w_P - \alpha\theta^1]h_P, \tag{35}$$

namely, the government would like to reduce w_P below the level where the left-hand side of (33) is zero if c_E is below the subsidy component of public employment at the

[17] An alternative terminology would be distributional marginal benefit and distributional marginal cost, as they reflect the direct marginal benefits and costs for society of redistributing resources from the taxpayers towards the targeted group, ceteris paribus, i.e., in the absence of any effects via the activity choices and the incentive compatibility constraints.

second-best optimum.

If self-selection constraint (14) does not bind so that $\mu_3 = 0$, an analogous sufficient condition for the second-best optimal wage in public employment to be high enough to satisfy $\phi u'(w_P h_P)/\gamma^0 - \lambda < 0$ becomes:

$$\Gamma_E^i = -c_E + (y_2^i - c_2^i) < [\alpha\theta^1 - w_P]h_P + (y_2^1 - c_2^1) = \Gamma_P, \quad i = 1, 2. \tag{36}$$

If condition (36) is satisfied, a higher wage in public employment would lead to more tax revenue. In this case, there is an incentive for the government to increase the number of type 0 individuals entering public employment in the first period and correspondingly decrease the number of individuals entering education.

Continuing with the policy rule for work hours in (33), we can see that the left-hand side reflects that a marginal increase in h_P affects the utility of type 0 individuals entering public employment by $w_P u'(w_P h_P) - v'(h_P) E_{f^i}(\xi \mid \xi \in \Omega_P^i)$ on average (with the welfare weight $\phi/\gamma^0 \geqslant 1$ attached to it), as well as reflects a social cost equal to $\lambda(w_P - \alpha\theta^1) > 0$. Thus, the direct marginal benefits and costs take more complex forms here than in (32). Note that if (i) type 0 individuals were paid their marginal product such that $w_P = \alpha\theta^1$, (ii) self-selection constraint (14) does not bind (such that $\mu_3 = 0$), and (iii) there are no behavioral responses in terms of activity choices among type 0 individuals, the right-hand side of (33) would be zero. In this case, the choice of h_P would be guided by an "average rule" for labor supply measured among those in public employment, where the consumption gain of an increase in h_P (weighted by ϕ/γ^0) is balanced against the value of lost leisure for the group as a whole.

Signing the right side of (33) is more difficult than signing the right side of (32). The reason is that the number of people in public employment can change in either direction as h_P increases. An increase in h_P generates more earned income but also a higher disutility of effort, making the overall effect on utility of remaining in the public employment state ambiguous. More specifically, if N_P^1 and N_P^2 increase in response to an increase in h_P, we can interpret the results in the same way as we did for an increase in w_P above, whereas the results would be the opposite if an increase in h_P leads to a decrease in N_P^1 and N_P^2, respectively.

4.3 Optimal education policy

By using μ_4 to denote the Lagrange multiplier attached to self-selection constraint (15), the policy rule for the education transfer is presented in proposition 3.

Proposition 3. *Under optimal income taxation, the optimal education transfer is charac-*

terized by the policy rule

$$\left[\frac{\phi}{\gamma^0}u'(c_E)-\lambda\right]\left(\sum_{i=1,2}N_E^i\right)+\sum_{i=1,2}H^{i'}(N_U^i)\frac{dN_U^i}{dc_E}$$
$$=-\lambda\left(\sum_{i=1,2}\left(\frac{dN_E^i}{dc_E}\Gamma_E^i\right)+\frac{d\left(N_P^1+N_P^2\right)}{dc_E}\Gamma_P-\frac{d\left(N_U^1+N_U^2\right)}{dc_E}(c_1^U+c_2^U)\right)+\frac{\mu_4}{\gamma^0}u'(c_E). \tag{37}$$

Proof See Appendix A.2. □

On the left-hand side of (37), $\phi u'(c_E)/\gamma^0-\lambda$ represents the difference between the direct marginal benefit of the education transfer - measured by the utility gain per recipient times the social welfare weight the government places on that utility gain - minus the direct marginal resource cost. In the absence of any behavioral response to this transfer among type 0 individuals, and in the absence of any incentive for type 1 to mimic type 0 (in which case $\mu_4=0$), c_E would be chosen such that $\phi u'(c_E)/\gamma^0-\lambda=0$. Thus, if the activity choices were independent of c_E, such a "first-best" decision rule would be second-best optimal as well in our model, since leisure is separable in terms of the utility function. This result is analogous to the Atkinson and Stiglitz (1976) theorem.

The second term on the left-hand side and the terms in parentheses on the right-hand side of (37) arise because the educational transfer affects the activity choices of individuals of type 0. Since $H^{i'}<0$ and $dN_U^i/dc_E<0$, the second term on the left-hand side is unambiguously positive and works to increase the education transfer beyond the level that balances the direct marginal benefits and marginal costs discussed above. The intuition is that the government attaches disutility to long-term unemployment, which provides an incentive to increase the number of type 0 individuals enrolling in education in the first period and correspondingly reduce the number of long-term unemployed. An alternative interpretation is that this component represents an individual-level social preference; namely, that each individual derives disutility from the severity of the problem of long-term unemployment.[18]

Thus, if (for some reason) the right-hand side of (37) were equal to zero, the results discussed so far would imply that the education transfer should be chosen so that $\phi u'(c_E)/\gamma^0-\lambda<0$, i.e., above the level that balances the direct marginal benefits and costs. An interesting question, therefore, is whether the second-best optimal education transfer could be even higher, such that the two terms on the left-hand side of (37) sum

[18] Such a preference may either reflect concern for the welfare of the (most likely) poorest group in society or a broader concern for the consequences that long-term unemployment may have (such as increased prevalence of crime). As long as the functions $H^i(\cdot)$ enter in an additively separable way, this change of assumption about the origin of these functions would neither affect the qualitative result presented in proposition 3 nor its interpretation.

to a negative number.

If self-selection constraint (15) does not bind, (i.e., $\mu_4 = 0$), a sufficient condition for the right-hand side of (37) to be negative becomes

$$\Gamma_E^i = -c_E + (y_2^i - c_2^i) > [\alpha\theta^1 - w_P]h_P + (y_2^1 - c_2^1) = \Gamma_P, \quad i = 1, 2. \tag{38}$$

Clearly, (38) is analogous to the condition under which it is welfare improving to reduce the wage in public employment below the level that balances direct marginal benefits and costs. This condition has an even stronger implication for the education transfer: if $\mu_4 = 0$ and (38) is satisfied, it is second-best optimal to provide c_E *in excess* of the level where the two terms on the left-hand side sum to zero, since an increase in educational attainment also leads to increased tax revenue. Since (38) is identical to (34), and if $\Gamma_E^2 > \Gamma_E^1$, we can also show, in the same way as above, that a sufficient condition for (38) to hold is $c_E < [w_P - \alpha\theta^1]h_P$. Therefore, the desire to collect tax revenue provides an incentive to simultaneously push up c_E and push down w_P relative to what would otherwise be optimal. This mechanism is counteracted when $\mu_4 > 0$, in which case the government can relax self-selection constraint (15) by lowering c_E. In this case, therefore, (38) is no longer a sufficient condition for educational provision beyond the level where $\phi u'(c_E)/\gamma^0 - \lambda = 0$.

4.4 Optimal unemployment benefits

Finally, we turn to the policy rules for unemployment benefits. An individual of type 0 who chooses unemployment in the first period will remain unemployed also in the second period, and the consumption stream of the unemployed will be $\{c_1^u, c_2^u\}$. The optimal unemployment benefits are characterized in Proposition 4.

Proposition 4. *Under optimal income taxation, the optimal unemployment benefits are characterized as follows (for* $t = 1, 2$*):*

$$\left[\frac{\phi}{\gamma^0}\beta^{t-1}u'(c_t^u) - \lambda\right] N_u + \sum_{i=1,2} H^{i'}(N_u^i)\frac{dN_u^i}{dc_t^u}$$
$$= \lambda\left(-\frac{dN_E^1}{dc_t^u}\Gamma_E^1 + \frac{d\left(N_u^1 + N_u^2\right)}{dc_t^u}(c_1^u + c_2^u)\right) \tag{39}$$

Proof See Appendix A.2. □

The direct marginal benefit per unemployed person, $\phi\beta^{t-1}u'(c_t^u)/\gamma^0$, and the direct marginal cost (in utility units), λ, appear in the brackets on the left side of the equation (39). We can also see that the second term on the left is clearly negative, since $H^{i'} < 0$

and $dN_U^i/dc_t^U > 0$ for $i = 1, 2$, meaning that the disutility society attaches to long-term unemployment works to reduce the unemployment benefit below the level at which the direct marginal benefit equals the direct marginal cost, ceteris paribus. Thus, if the right-hand side of the equation (39) were equal to zero (for whatever reason), the unemployment benefit would satisfy the condition $\phi\beta^{t-1}u'(c_t^U)/\gamma^0 - \lambda > 0$.

To interpret the right-hand side, note that only two of the activity choices, unemployment and education leading to productivity θ^1, are directly affected by a change in unemployment benefits. Also note that $dN_E^1/dc_t^U = -dN_U^1/dc_t^U$. Therefore, the sign of the right-hand side of the equation (39) is positive if, and only if, $\Gamma_E^1 + c_1^U + c_2^U > 0$, i.e,

$$c_1^U + c_2^U > c_E - (y_2^1 - c_2^1). \tag{40}$$

Inequality (40) means that public expenditure per unemployed individual, measured over the two periods, exceeds net public expenditure (public expenditure minus tax revenue) per type 0 individual who enrolls in education and realizes productivity θ^1 in the second period. If this condition is satisfied, it works to further reduce the unemployment benefit below the level where the left-hand side of (39) equals zero.[19] Again, the intuition is that public funds are costly, and that a decrease in the number of unemployed (with a corresponding increase in educational attainment) leads to increased tax revenue. The opposite conclusion would arise in the (less likely) scenario where inequality (40) is reversed, in which case there is an incentive to increase unemployment benefits to avoid pressure on the public funds.

5 Numerical simulations

To gain further insights into the policy implications of long-term unemployment, we will now turn to numerical simulations. We begin by describing the key components of the simulations: (i) the distribution of innate ability, (ii) the distribution of effort costs, and (iii) the functional form of the utility function and the parameters of human capital production.

5.1 The distribution of innate abilities

The distribution of market wage rates used to approximate the skill distribution is based on Swedish wage data from 2016, where we assume that θ^1 represents the 10th wage

[19] As we indicated above, we have neglected the possibility that a true type 1 individual could mimic the unemployment choice made by some of the type zero agents in order to benefit from the redistribution to the unemployed. If we were to add such a self-selection constraint (and in the unlikely case that the constraint binds), it would further contribute to the "underprovision result," i.e, that the policy rule for unemployment benefits satisfies $\phi\beta^{i-1}u'(b_i)/\gamma^0 - \lambda > 0$.

percentile and θ^2 the 80th percentile of the hourly wage distribution.[20] This implies that $\theta^1/\theta^2 \approx 0.58$.

We are agnostic about the latent productivity of type 0 by assuming $\pi^1 = \pi^2 = 0.5$, i.e., an equal fraction of latent low-skilled and high-skilled agents of type 0 in the economy. In the baseline parameterization, $\alpha = 0.9$, meaning that individuals of type 0 have a market productivity equal to 90% of the productivity of conventional type 1 agents in the first period.

Type 0 agents make up 10% of the population in the baseline simulation, which means that $\gamma^0 = 0.1$.[21] The remaining share of the population is divided into 40% low-productivity agents (conventional type 1, with labor productivity θ^1) and 60% high-productivity agents (conventional type 2, with labor productivity θ^2). Thus, $\gamma^1 = 0.4 \times (1-\gamma^0) = 0.36$ and $\gamma^2 = 0.6 \times (1-\gamma^0) = 0.54$.

5.2 The distribution of effort costs

We assume that the effort cost, ξ, for type 0 agents is independent of θ and follows a generalized Pareto distribution with location, shape, and scale parameters all equal to 1. With these parameter choices, the density reduces to

$$f(x) = x^{-2}, \quad x \geqslant 1.$$

Thus, the effort cost has support on $[1, \infty)$ and is heavy-tailed, so that large effort costs occur with non-negligible probability, affecting the occupational choices of individuals.

5.3 Utility function

Following, e.g., Conesa et al. (2009) and Bastani et al. (2013), consumption utility is given by

$$u(c) = \frac{c^{1-\eta} - 1}{1-\eta}, \tag{41}$$

while the function describing the disutility of labor supply/effort takes the following form:

$$v(h) = \zeta \frac{h^k}{k}. \tag{42}$$

We set $k = 3$ corresponding to a constant consumption (Frisch) labor supply elasticity equal to $1/(k-1) = 0.5$, which is close to the central estimate of 0.4 reported in the meta-study of Whalen and Reichling (2017), and set $\eta = 3$ corresponding to an elasticity

[20] The data are administered by the Swedish National Mediation Office and include monthly wages, expressed in full-time equivalents, for all workers in the public sector and about 50 percent of all workers in the private sector.

[21] According to Statistics Sweden, about 13% of the population aged 20-64 (expressed in full-time equivalents) was supported by transfers or social assistance in Sweden in 2020.

of intertemporal substitution of 1/3, consistent with recent studies such as Best et al. (2019) and Ring (2024). The scaling parameter ζ is calibrated so that the labor supply among low-skilled workers is roughly equal to 1 in the baseline parameterization.

We also assume that the hours in public employment cannot exceed full-time ($h_P \leqslant 1$). In other words, the government does not encourage overtime among people in public employment, which we believe realistically captures the government's adherence to prevailing employment protection regulations. Regarding educational effort, we have chosen $e^1 = e^2 = \underline{e} = e = 0.75$, corresponding to 75% of the calibrated value of 1 for the labor supply among low-skilled workers.

5.4 Social objective function and baseline parameterization

In the simulations, we abstract from the (possible) disutility that the government attaches to long-term unemployment by omitting the functions $H^i(\cdot), i = 1, 2$ in equation (16).[22] We also focus on a Utilitarian planner ($\phi = \gamma^0$) as our baseline. Our baseline parameterization is summarized in Table 1.

Table 1: Parameters in baseline calibration

Parameter	Description	Value
θ^1	Productivity of type 1 agents	1.42
θ^2	Productivity of type 2 agents	2.43
α	Productivity type 0 agents (as a share of θ^1)	0.90
γ^0	Share of type 0 agents	0.10
ϕ	Welfare weight of type 0 agents	0.10
γ^1	Share of type 1 agents	0.36
γ^2	Share of type 2 agents	0.54
e	Effort requirement in education	0.75
π^1	Share of type 0 agents with productivity θ^1	0.50
π^2	Share of type 0 agents with productivity θ^2	0.50
η	Coefficient of relative risk aversion	3
k	Labor supply elasticity parameter	3

[22] Simulations with these functions were included in an earlier version of our paper, Aronsson et al. (2021), where we showed, perhaps unsurprisingly, that aversion to long-term unemployment leads to less generous unemployment benefits and thus to lower long-term unemployment relative to the baseline.

6 Quantitative results

6.1 Baseline results

The results for the baseline parameterization are presented in Table 2.[23] With a slight abuse of notation, we have expressed the number of individuals enrolled in each of the different occupational states as a percentage of the total number of Type 0 individuals of each latent productivity type. For example, N_E^1 is measured by dividing equation (22) by $\gamma^0\pi^1$ such that $N_E^1 = F^1(\xi_{E-U}) - F^1(\xi_{P-E})$, and similarly for the number of individuals entering the other occupational states in equations (22)–(24).[24]

Table 2: Baseline results

Tax rates		Occupation shares		Pre/post-tax income		Other	
$\tau_{y,1}^1$	0.35	N_P^1	0.69	y_1^1	1.52	c_1^U	1.29
$\tau_{y,1}^2$	0.00	N_E^1	0.17	y_2^1	1.55	c_2^U	1.31
$\tau_{y,2}^1$	0.37	N_U^1	0.14	y_1^2	2.90	c_E	1.27
$\tau_{y,2}^2$	0.00	N_P^2	0.36	y_2^2	2.98	w_P	1.42
T_1^1	-0.33	N_E^2	0.54	c_1^1	2.01	h_P	1.00
T_2^1	-0.26	N_U^2	0.11	c_2^1	1.95	h_1^1	1.07
T_1^2	0.11			c_1^2	2.58	h_2^1	1.09
T_2^2	0.15			c_2^2	2.53	h_1^2	1.19
						h_2^2	1.23

[23]The model is set up numerically using Julia and solved using the state-of-the-art constrained optimization solver KNITRO developed by Artelys Inc. In this way, we follow previous work that has used KNITRO to solve numerically challenging optimal income tax problems, see, for example, Bastani et al. (2013, 2020).

[24]In the theory part, we assumed a logical ordering of the integration constraints such that individuals choose activities in the following order (depending on the effort cost): public employment, education, and unemployment. In the simulations, we impose these constraints to discipline the solution procedure, but then verify that they are not binding in the social optimum.

Figure 2: Distribution of effort costs and activity choice thresholds

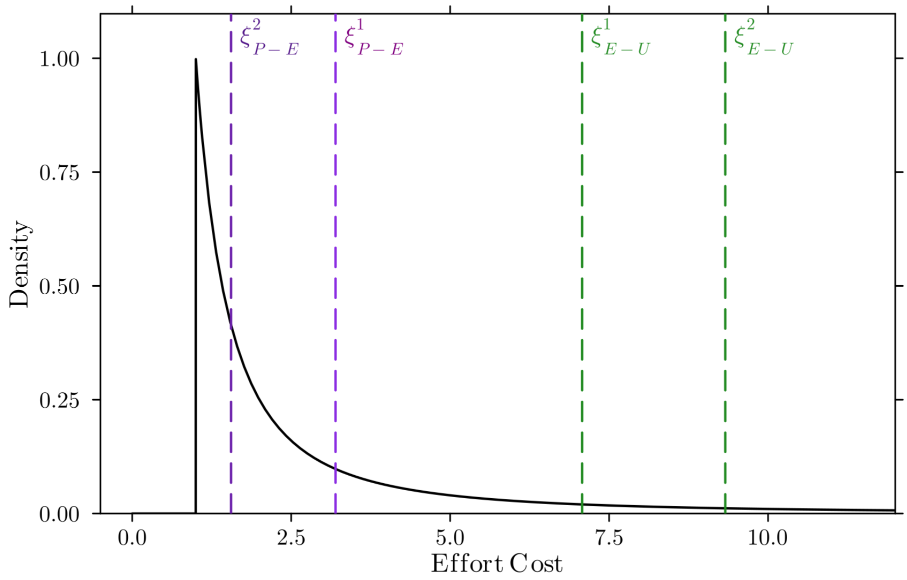


Figure 2 illustrates the distribution of effort costs and the thresholds determining activity choices for type 0 agents. The figure shows how the effort cost thresholds ξ^1_{P-E}, ξ^1_{E-U}, ξ^2_{P-E}, and ξ^2_{E-U} divide the population into different occupational choices. For type 0 agents with latent productivity θ^1, those with effort costs below ξ^1_{P-E} choose public employment, those between ξ^1_{P-E} and ξ^1_{E-U} choose education, and those above ξ^1_{E-U} choose unemployment. Similarly, for type 0 agents with latent productivity θ^2, the thresholds ξ^2_{P-E} and ξ^2_{E-U} determine their occupational choices. The figure helps visualize why we observe the occupational shares reported in Table 2.

Let us begin with the occupational choices. Among type 0 individuals with latent low skills, we can see that 69% choose public employment, 17% choose education, and 14% choose unemployment. The corresponding numbers for type 0 individuals with latent high skills are 36% in public employment, 54% in education, and 11% in unemployment. Thus, whereas the dominating activity choice among the latent low-skilled is public employment, the dominating activity choice among the latent high-skilled is education. In addition, it is (slightly) more common among the latent low-skilled to end up in long-term unemployment.

In accordance with the theoretical results, the marginal income tax rates implemented for the high-skilled are zero, while the marginal income tax rates implemented for the low-skilled are positive in both periods. These distortions are necessary in order to support redistribution in favor of the low-skilled and reflect binding incentive compatibility constraints. The average tax rates are negative for the low-skilled (meaning that part of their disposable income originates from governmental transfers) and positive for the high-skilled. The numbers in Table 2 indicate a substantial redistribution from high-skilled to low-skilled workers.

The labor supply patterns show that high-skilled workers supply more labor than low-skilled workers in both periods. This result is consistent with the lower marginal tax rates implemented for the high-skilled, and ensures that those with higher earnings ability also earn higher incomes than those with lower earnings ability.

Finally, turning to the public expenditure programs targeting type 0 individuals, we can see that the wage in public employment ($w_P = 1.42$) is equal to the productivity of type 1 workers, while the hours requirement ($h_P = 1.00$) corresponds to full-time employment. The education transfer ($c_E = 1.27$) is set such that education becomes attractive for those with moderate effort costs, while the unemployment benefits (1.29 and 1.31 in periods 1 and 2, respectively) provide a basic safety net for those ending up in long-term unemployment.

6.2 Numerical comparative statics

We will now use the numerical model to perform comparative statics analyses. Our investigation focuses on three key dimensions: the government's preference for redistribution, the dispersion of the productivity distribution, and the proportion of type 0 agents in the economy.[25] These three dimensions are central determinants of the optimal redistribution policy in general and the policies aimed at targeting the activity choices among type 0 agents in particular.[26]

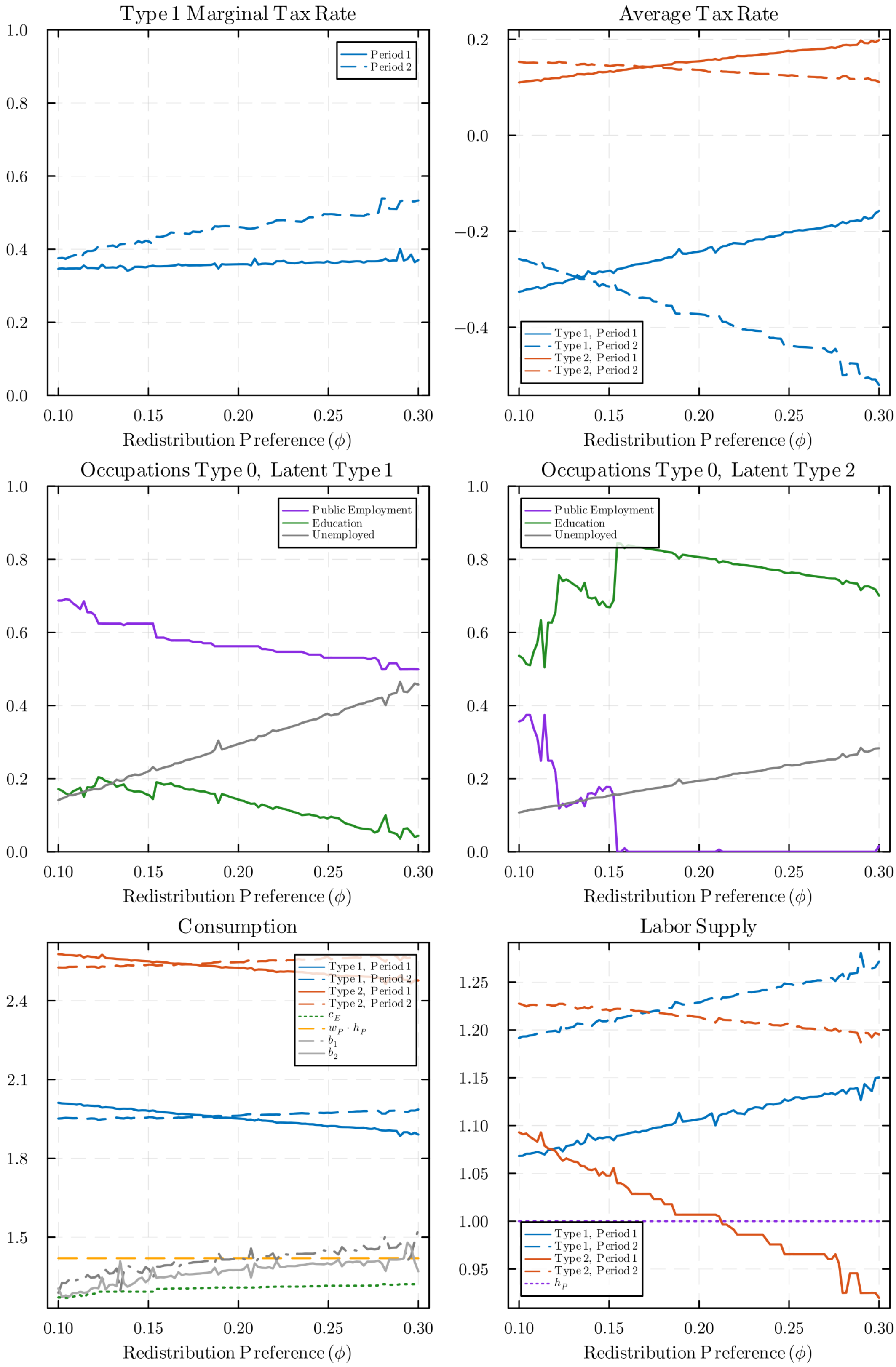
Government's taste for redistribution We begin by exploring the role of the government's taste for redistribution by gradually increasing (relative to the utilitarian benchmark) the social welfare weight assigned to type 0 agents, from $\phi = 0.1$ (the benchmark value) to $\phi = 0.3$. The results are shown in Figure 3.[27]

[25] In these exercises, we vary one such parameter at a time while holding the others constant, consistent with traditional comparative statics methodology.

[26] As explained above, variation in the share of type 0 individuals can also be interpreted in terms of changes in migration flows, while increased wage dispersion has a natural interpretation in terms of skill-biased technological change.

[27] The jaggedness in the graphs stems from numerical integration in the optimization process. However, as shown in Figure 6, the underlying utility components evolve relatively smoothly across the comparative statics parameters, indicating that the observed irregularities are purely numerical artifacts rather than substantive features of the model.

Figure 3: Increase in the government's taste for redistribution (baseline is $\phi = 0.1$)



As the social welfare weight attached to type 0 agents increases, the marginal tax rate implemented for type 1 workers increases in the second period, with a corresponding decrease in the average tax rate and increase in consumption. This is explained by the fact that type 0 agents choosing public employment or education in the first period become type 1 workers in the second period. Note also that the unemployment benefit increases in both periods, reflecting the desire to increasingly redistribute in favor of the long-term unemployed. Thus, there are two distinct channels of redistribution towards type 0: (i) direct transfers to ordinary type 1 workers in period 2, combined with a higher marginal tax rate to discourage mimicking, and (ii) increased unemployment benefits. Whereas the first channel targets those entering the regular labor market in the second period, the second channel targets the long-term unemployed. The increased redistribution in favor of type 0 goes through both these channels as ϕ increases.

For type 0 agents with latent productivity θ^2, there is a shift from public employment to education. Inducing these agents to attend education is costly because they do not produce in the first period. However, if ϕ is high, it is, nevertheless, socially desirable to push a larger fraction of these agents to realize a higher market productivity in period 2.

Increasing the dispersion of the productivity distribution We analyze the effects of a mean-preserving spread in productivity by varying θ^1 and θ^2 while keeping their mean constant. Define

$$m = \frac{\theta^1 + \theta^2}{2}, \quad d = \frac{\theta^2 - \theta^1}{2},$$

and set

$$\tilde{\theta}^1 = m - d\,\lambda, \quad \tilde{\theta}^2 = m + d\,\lambda,$$

with λ varying from 0.75 to 1.25, thereby altering the ratio θ^2/θ^1 without changing the average productivity. The simulation results are shown in Figure 4.

Figure 4: Change in the spread of the productivity distribution θ^2/θ^1 (baseline is $\lambda = 1$)

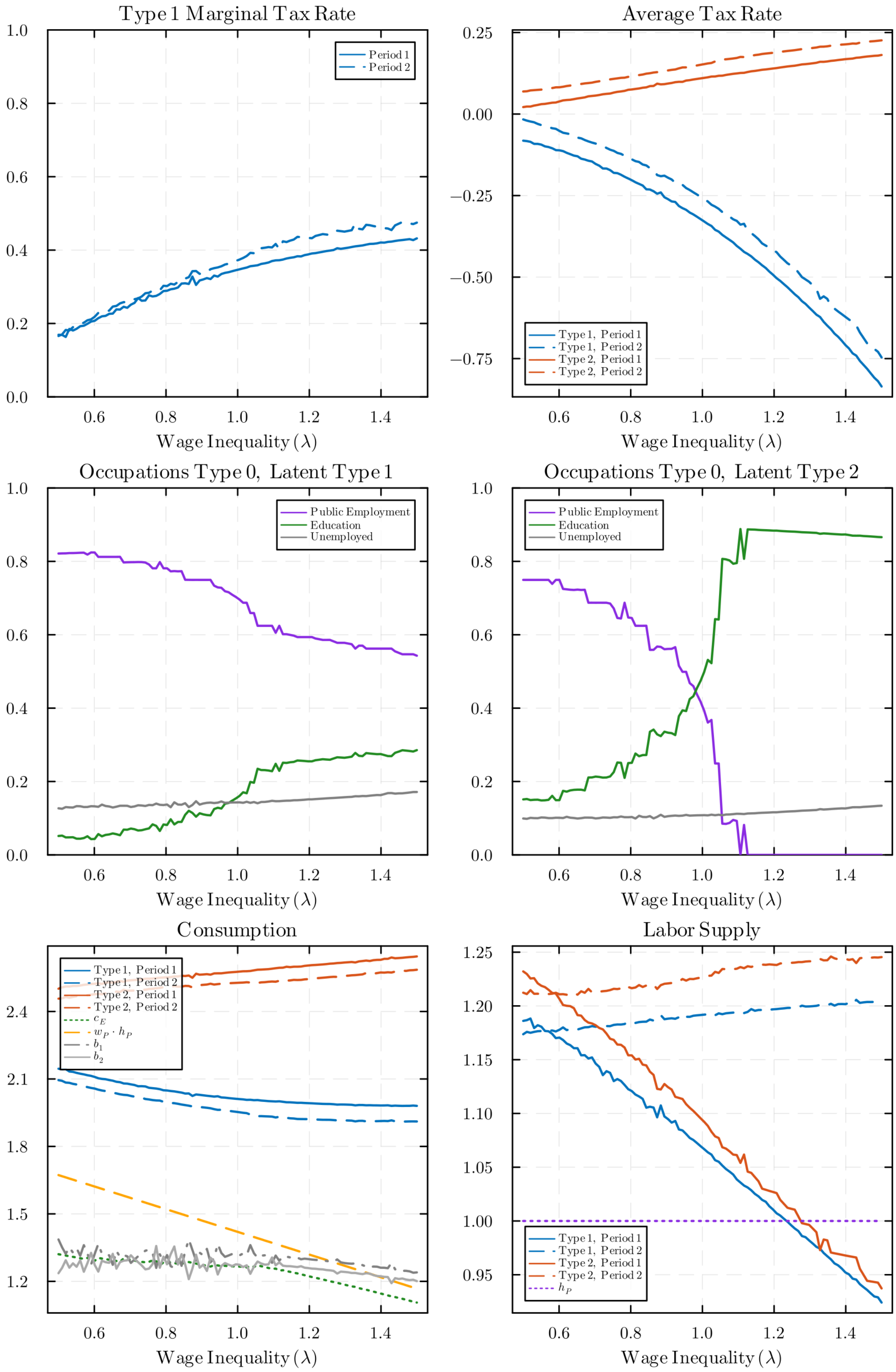
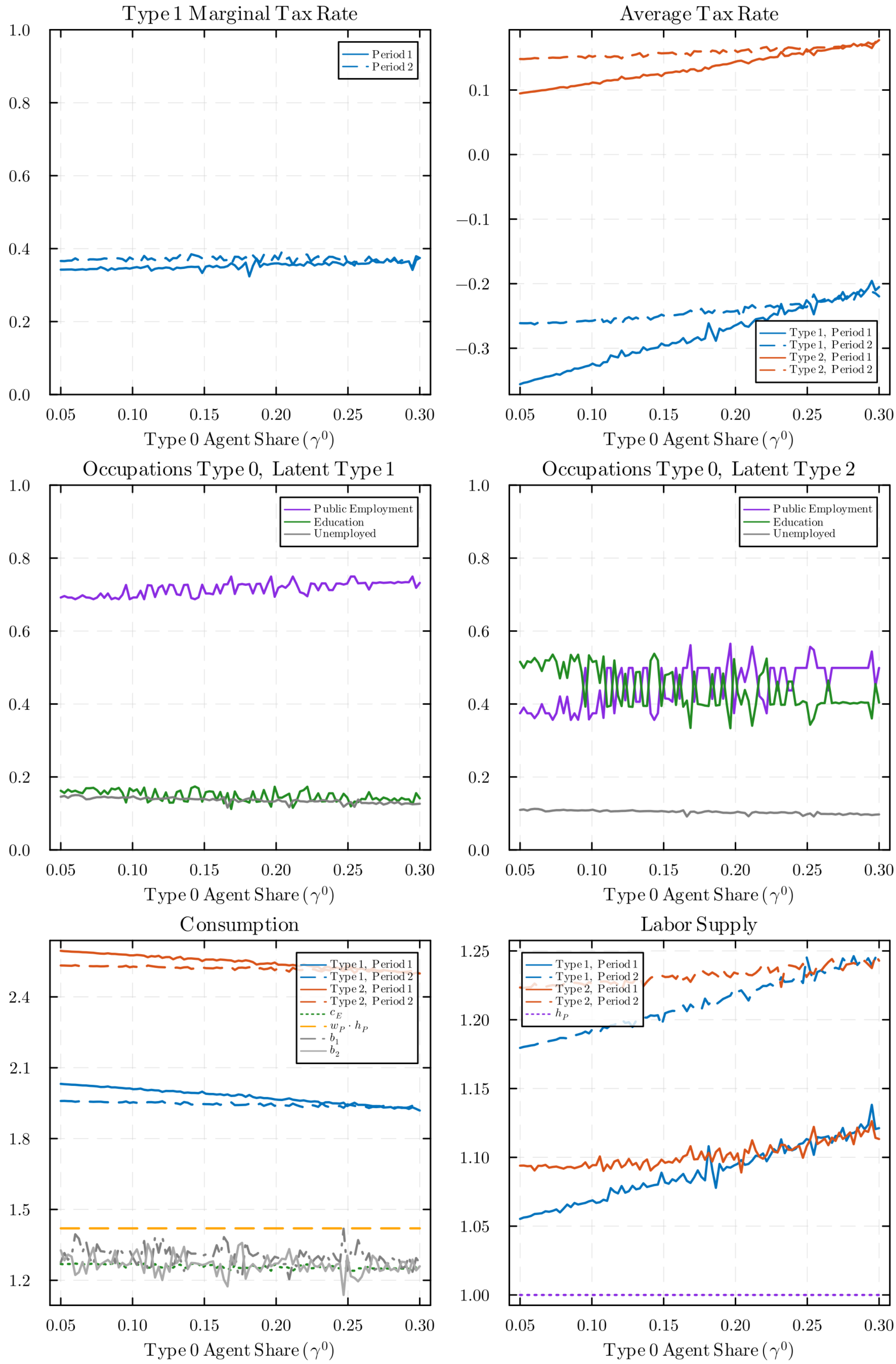


Figure 4 shows that increased productivity disparity leads to more redistribution in favor of type 1 workers. This result was not unexpected, given the concavity of the individual utility functions. As the skill level of type 1 workers decreases and the skill level of type 2 workers increases, the implicit social welfare weight of the former group increases with a corresponding decrease in the social welfare weight of the latter.

A mean-preserving spread of the productivity distribution affects occupational choices among type 0 agents in a way similar to a stronger governmental preference for redistribution. As the skill gap increases, the number of agents in public employment decreases among both latent skill types with corresponding increases in educational attainment. We can also see that higher skill differentials imply more long-term unemployment at the social optimum. This is because public employment is now a less attractive choice than before.

Share of type 0 agents in the economy Finally, we investigate the effects of changing the proportion of type 0 agents, γ^0, considering values from 0.05 to 0.3, recalling that the baseline value of γ^0 is 0.1. Consistent with our benchmark, we assume an equal distribution of latent productivity among these agents, setting $\pi^1 = \pi^2 = 0.5$. The corresponding simulation results are presented in Figure 5.

Figure 5: Changes in the share of type 0 agents (baseline is $\gamma^0 = 0.1$)

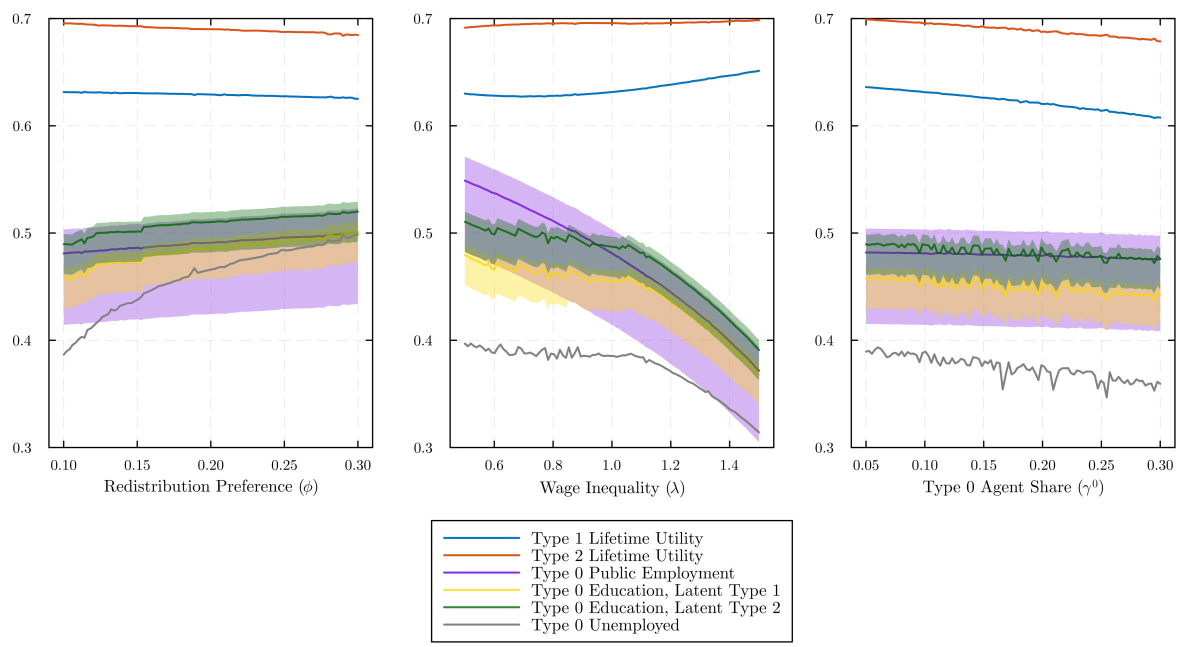


Note first that increases in γ^0 has little effect on the occupational choices among type 0 agents. The general pattern is that an increase in the share of type 0 agents leads to a higher fiscal burden on type 1 and type 2 workers, reflected in higher average tax rates and declining consumption. Labor supply is also increasing.

6.3 Welfare Analysis

To better understand the welfare implications of the comparative statics analyses, we decompose the social welfare function into its key components. Figure 6 shows how these different components of social welfare respond to changes in our key parameters.

Figure 6: Welfare components under different parameterizations



Note: For Public Employment and Education, solid lines represent welfare at the median effort cost $\xi = 2.00$. Shaded ribbons indicate the welfare range corresponding to the 25th (upper edge, $\xi = 1.33$) and 75th (lower edge, $\xi = 4.00$) percentiles of the ξ distribution. All entries refer to life-time utility.

The welfare decomposition provides several important insights. First, ordinary type 1 and type 2 workers are generally better off than their type 0 counterparts. Second, an increase in the welfare weight attached to type 0 workers contributes to reduce this gap. We can also see that a mean preserving increase in the wage dispersion makes ordinary type 1 workers better off (through increased redistribution), type 0 workers generally worse off, and leaves the welfare of ordinary type 2 workers relatively unaffected. Finally, an increase in the share of type 0 workers, as reflected by the third column, makes all workers worse off, perhaps most so for those in unemployment.

7 Concluding remarks

This paper is the first to simultaneously analyze the optimal tax policies, transfer mechanisms, and public spending programs needed to address the challenges posed by individuals with limited labor market attachment. In our framework, long-term unemployment occurs when certain individuals lack the productivity to secure employment in the regular labor market. These individuals, referred to as "type 0" agents, may belong to either latent low-skill or high-skill categories, but are unable to translate these skills into market productivity without the necessary education or training.

We would like to highlight four main conclusions. First, the policy rules for marginal income taxation are broadly consistent with those derived from standard models of optimal taxation. This consistency arises because the activity choices of type 0 agents are largely unaffected by marginal tax policy. However, the endogeneity of the skill distribution resulting from the influx of individuals into the labor market following successful integration and the increased tax demands resulting from long-term unemployment have an impact on both marginal and average tax rates.

Second, the optimal design of the education program, the public employment program, and unemployment benefits underscores the central role played by the scarcity of public resources, which in turn created a "tax revenue motive" behind each of these instruments. This revenue motive serves as a key determinant of whether education and public employment should be expanded or contracted relative to standard policy rules that do not take into account the activity choices of type 0 agents. Under reasonable assumptions, we find that education should be expanded relative to such a policy rule, while public employment tends to be contracted.

Third, the dilemma of long-term unemployment significantly limits the scope for redistribution to individuals who are not actively engaged in skill development or validation. The rationale behind the second and third conclusions is the government's dual objective of redistributing resources in favor of the low-skilled while mitigating the challenges posed by long-term unemployment.

Fourth, using a calibrated numerical model, we show that the government's propensity to redistribute and the dispersion of the skill distribution play a crucial role in shaping outcomes for type 0 agents. In particular, increasing the dispersion of productivity can significantly increase the optimal share of long-term unemployed in a second-best scenario.

Our study represents a first attempt to integrate the complexities of long-term unemployment into the theory of optimal redistributive taxation and public spending. Future research can explore extensions such as incorporating unemployment cultures or norms that may be partially inherited, possibly through the use of an overlapping generations model. Integrating social aversion to long-term unemployment with poverty aversion and poverty alleviation would be another interesting direction for future research.

References

Aronsson, T., Bastani, S., and Tayibov, K. (2021). Social exclusion and optimal redistribution. Working Paper 9448, CESifo.

Aronsson, T. and Johansson-Stenman, O. (2021). A note on optimal taxation, status consumption, and unemployment. *Journal of Public Economics*, 200:104458.

Aronsson, T. and Micheletto, L. (2021). Optimal redistributive income taxation and efficiency wages. *Scandinavian Journal of Economics*, 123(1):3–32.

Aronsson, T. and Sjögren, T. (2004). Efficient taxation, wage bargaining and policy coordination. *Journal of Public Economics*, 88:2711–2725.

Aronsson, T. and Sjögren, T. (2016). Quasi-hyperbolic discounting, paternalism and optimal mixed taxation. *Mathematical Social Sciences*, 84:24–36.

Atkinson, A. B. (1998). Social exclusion, poverty and unemployment. In Hills, J., editor, *Exclusion, Employment and Opportunity*, pages 1–20. Centre for Analysis of Social Exclusion (CASE), London School of Economics and Political Science, London.

Atkinson, A. B. and Stiglitz, J. E. (1976). The design of tax structure: Direct versus indirect taxation. *Journal of Public Economics*, 6:55–75.

Bastani, S., Blomquist, S., and Micheletto, L. (2013). The welfare-gains of age related optimal income taxation. *International Economic Review*, 54(4):1219–1249.

Bastani, S., Blomquist, S., and Micheletto, L. (2020). Child care subsidies, quality and optimal income taxation. *American Economic Journal: Economic Policy*, 12(1):1–37.

Bastani, S., Blumkin, T., and Micheletto, L. (2015). Optimal wage redistribution in the presence of adverse selection in the labor market. *Journal of Public Economics*, 131:41–57.

Bastani, S., Blumkin, T., and Micheletto, L. (2025). Optimal redistribution and education signaling. *International Economic Review*.

Best, M. C., Cloyne, J. S., Ilzetzki, E., and Kleven, H. J. (2019). Estimating the elasticity of intertemporal substitution using mortgage notches. *The Review of Economic Studies*, 87(2):656–690.

Blumkin, T., Margalioth, Y., and Sadka, E. (2007). Anti-discrimination rules versus income taxation in the pursuit of horizontal equity. *Journal of Public Economics*, 91:1167–1176.

Boadway, R., Marchand, M., and Pestieau, P. (1991). Optimal linear income taxation in models with occupational choice. *Journal of Public Economics*, 46(2):133–162.

Bovenberg, A. L. and Jacobs, B. (2005). Redistribution and education subsidies are siamese twins. *Journal of Public Economics*, 89:2005–2035.

Bradley, S., Crouchley, R., and Oskrochi, R. (2003). Social exclusion and labor market transitions: A multi-state multi-spell analysis using the BHPS. *Labour Economics*, 10:659–679.

Brett, C. and Weymark, J. A. (2019). Optimal nonlinear taxation of income and savings without commitment. *Journal of Public Economic Theory*, 21(1):5–43.

Clark, A. E. and Oswald, A. J. (1994). Unhappiness and unemployment. *The Economic Journal*, 104(424):648–659.

Colas, M., Findeisen, S., and Sachs, D. (2021). Optimal need-based financial aid. *Journal of Political Economy*, 129(2):492–533.

Conesa, J. C., Kitao, S., and Krueger, D. (2009). Taxing capital? not a bad idea after all! *American Economic Review*, 99(1):25–48.

Costinot, A. and Werning, I. (2023). Robots, trade, and luddism: A sufficient statistic approach to optimal technology regulation. *Review of Economic Studies*, 90(5):2261–2291.

Craig, A. C. (2023). Optimal income taxation with spillovers from employer learning. *American Economic Journal: Economic Policy*, 15(2):82–125.

Da Costa, C. E. and Santos, M. R. (2018). Age-dependent taxes with endogenous human capital formation. *International Economic Review*, 59:785–823.

Diamond, P. A. (1980). Income taxation with fixed hours of work. *Journal of Public Economics*, 13:101–110.

Findeisen, S. and Sachs, D. (2016). Education and optimal dynamic taxation: The role of income-contingent student loans. *Journal of Public Economics*, 138:1–21.

Guerreiro, J., Rebelo, S., and Teles, P. (2022). Should robots be taxed? *Review of Economic Studies*, 89(1):279–311.

Hammond, P. J. (1979). Straightforward individual incentive compatibility in large economies. *Review of Economic Studies*, 46(2):263–282.

Henry de Frahan, L. and Maniquet, F. (2021). Preference responsibility versus poverty reduction in the taxation of labor incomes. *Journal of Public Economics*, 197:104386.

Hungerbühler, M., Lehmann, E., Parmentier, A., and Van der Linden, B. (2006). Optimal redistributive taxation in a search equilibrium model. *The Review of Economic Studies*, 73(3):743–767.

Kanbur, R., Paukkeri, T., Pirttilä, J., and Tuomala, M. (2018). Optimal taxation and public provision for poverty reduction. *International Tax and Public Finance*, 25:64–98.

Kieselbach, T. (2003). Long-term unemployment among young people: The risk of social exclusion. *American Journal of Community Psychology*, 32:69–76.

Lehmann, E., Parmentier, A., and Van Der Linden, B. (2011). Optimal income taxation with endogenous participation and search unemployment. *Journal of Public Economics*, 95:1523–1537.

Lockwood, B. B., Nathanson, C. G., and Weyl, E. G. (2016). Taxation and the allocation of talent. *Journal of Political Economy*, 125:1635–1682.

Makris, M. and Pavan, A. (2021). Taxation under learning-by-doing. *Journal of Political Economy*, 129:1878–1944.

Maldonado, D. (2008). Education policies and optimal taxation. *International Tax and Public Finance*, 15:131–143.

Maniquet, F. and Neumann, D. (2021). Well-being, poverty, and labor income taxation: Theory and application to europe and the united states. *American Economic Journal: Microeconomics*, 13(2):276–310.

Marceau, N. and Boadway, R. (1994). Wage legislation and unemployment insurance as instruments for redistribution. *Scandinavian Journal of Economics*, 96:67–81.

Mirrlees, J. A. (1971). An exploration in the theory of optimum income taxation. *Review of Economic Studies*, 38:175–208.

Mörk, E., Ottosson, L., and Vikman, U. (2021). To work or not to work? effects of temporary public employment on future employment and benefits. Working Paper 2021:12, IFAU.

Pirttilä, J. and Tuomala, M. (2004). Poverty alleviation and tax policy. *European Economic Review*, 48:1075–1090.

Pohlan, L. (2019). Unemployment and social exclusion. *Journal of Economic Behavior and Organization*, 164:273–299.

Ring, M. A. K. (2024). Wealth taxation and household saving: Evidence from assessment discontinuities in norway. *The Review of Economic Studies*. rdae100.

Rothschild, C. and Scheuer, F. (2013). Redistributive taxation in the roy model. *Quarterly Journal of Economics*, 128:623–668.

Rothschild, C. and Scheuer, F. (2016). Optimal taxation with rent-seeking. *Review of Economic Studies*, 83:1225–1262.

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

Saez, E. and Stantcheva, S. (2016). Generalized social marginal welfare weights for optimal tax theory. *American Economic Review*, 106(1):24–45.

Sen, A. (2000). Social exclusion: Concept, application, and scrutiny. Social Development Papers 1, Office of Environment and Development, Asian Development Bank.

Stantcheva, S. (2014). Optimal income taxation with adverse selection in the labour market. *The Review of Economic Studies*, 81(3):1296–1329.

Stantcheva, S. (2017). Optimal taxation and human capital policies over the life cycle. *Journal of Political Economy*, 125(6):1931–1990.

Stern, N. H. (1982). Optimum taxation with errors in administration. *Journal of Public Economics*, 17:181–211.

Stiglitz, J. E. (1982). Self-selection and pareto efficient taxation. *Journal of Public Economics*, 17:213–240.

Sztutman, A. (2024). Dynamic job market signaling and optimal taxation. *Working Paper, Tepper School of Business, Carnegie Mellon University.*

Thuemmel, U. (2023). Optimal taxation of robots. *Journal of the European Economic Association*, 21(3):1154–1190.

Whalen, C. and Reichling, F. (2017). Estimates of the frisch elasticity of labor supply: A review. *Eastern Economic Journal*, 43(1):37–42.

Winkelmann, L. and Winkelmann, R. (1998). Why are the unemployed so unhappy? evidence from panel data. *Economica*, 65(257):1–15.

A Proofs of Propositions

The Lagrangean of the social decison-problem cn be written as

$$
\begin{aligned}
\mathcal{L} =& \phi\Big(\pi^1 \left(\int_0^{\xi^1_{P-E}} V_P(\theta^1,\xi)f(\xi)d\xi + \int_{\xi^1_{P-E}}^{\xi^1_{E-U}} V_E(\theta^1,\xi)f(\xi)d\xi + \int_{\xi^1_{E-U}}^{\infty} V_U(\theta^1,\xi)f(\xi)d\xi\right) \\
&+ \pi^2\Big(\int_0^{\xi^2_{P-E}} V_P(\theta^2,\xi)f(\xi)d\xi \\
&+ \int_{\xi^2_{P-E}}^{\xi^2_{E-U}} V_{E^1}(\theta^2,\xi)f(\xi)d\xi + \int_{\xi^2_{E-U}}^{\infty} V_U(\theta^2,\xi)f(\xi)d\xi\Big)\Big) \\
&+ H^1(\gamma^0\pi^1[1-F(\xi^1_{E-U})]) + H^2(\gamma^0\pi^2[1-F(\xi^2_{E-U})]) \\
&+ \gamma^1[u(c^1_1) - v(y^1_1/\theta^1) + \beta\left(u(c^1_2) - v(y^1_2/\theta^1)\right)] \\
&+ \gamma^2[u(c^2_1) - v(y^2_1/\theta^2) + \beta\left(u(c^2_2) - v(y^2_2/\theta^2)\right)] \\
&+ \mu^1\Big(u(c^2_1) - v(y^2_1/\theta^2) - [u(c^1_1) - v(y^1_1/\theta^2)]\Big) \\
&+ \mu^2\Big(u(c^2_2) - v(y^2_2/\theta^2) - [u(c^1_2) - v(y^1_2/\theta^2)]\Big) \\
&+ \mu_3\Big(u(c^1_1) - v(y^1_1/\theta^1) - u(w_P h_P) + v_P(h_P)\Big) \\
&+ \mu_4\Big(u(c^1_1) - v(y^1_1/\theta^1) - u(c_E) + v_E(\underline{e})\Big) \\
&+ \lambda\Big(\gamma^1\Big[(y^1_1 - c^1_1) + (y^1_2 - c^1_2)\Big] + \gamma^2\Big[(y^2_1 - c^2_1) + (y^2_2 - c^2_2)\Big] \\
&+ \gamma^0\left(\pi^1[F(\xi^1_{E-U}) - F(\xi^1_{P-E})]\right)\left(-c_E + (y^1_2 - c^1_2)\right) \\
&+ \gamma^0\,\pi^2\left(F(\xi^2_{E-U}) - F(\xi^2_{P-E})\right)\left(-c_E + (y^2_2 - c^2_2)\right) \\
&+ \gamma^0\left(\pi^1 F(\xi^1_{P-E}) + \pi^2 F(\xi^2_{P-E})\right)\left([\alpha\theta^1 - w_P]h_P + (y^1_2 - c^1_2)\right) \\
&- \gamma^0\left(\pi^1[1-F^{(}\xi^1_{E-U})] + \pi^2[1-F(\xi^2_{E-U})]\right)(b_1+b_2) - M\Big).
\end{aligned} \tag{A1}
$$

In (A1), we have used the following notation:

$$V_P(\theta^i,\xi) = u(w_P h_P) - \xi v_P(h_P) + \beta\left(u(c^1_2) - v(y^1_2/\theta^1)\right), \quad i = 1,2 \tag{A2}$$

$$V_U(\theta^i,\xi) = u(b_1) + \beta u(b_2), \quad i = 1,2 \tag{A3}$$

$$V_E(\theta^1,\xi) = u(c_E) - \xi v(e^1) + \beta\left(u(c^1_2) - v(y^1_2/\theta^1)\right) \tag{A4}$$

$$V_{E^2}(\theta^2,\xi) = u(c_E) - \xi v(e^2) + \beta\left(u(c^2_2) - v(y^2_2/\theta^2)\right). \tag{A5}$$

The limits of integration for ξ in (A1) are given by

$$\xi^1_{P-E} = \frac{u(w_P h_P) - u(c_E)}{v(h_P) - v(e^1)} > 0. \tag{A6}$$

$$\xi^2_{P-E} = \frac{[u(w_P h_P) - u(c_E)] + \beta\left([u(c^1_2) - u(c^2_2)] - [v(y^1_2/\theta^1) - v(y^2_2/\theta^2)]\right)}{v(h_P) - v(e^2)}. \tag{A7}$$

$$\xi^i_{E-U} = \frac{1}{v(e^i)}\left([u(c_E) - u(c^U_1)] + \beta\left([u(c^i_2) - v(y^i_2/\theta^i)] - u(c^U_2)\right)\right), i = 1, 2. \tag{A8}$$

A.1 Derivation of optimal labor distortions (Proposition 1)

First period The social first-order conditions for labor supply and consumption can be written as period 1 can be written as:

$$y^1_1 : (\gamma^1 + \mu_3 + \mu_4)v'(y^1_1/\theta^1)/\theta^1 - \mu_1 v'(y^1_1/\theta^2)/\theta^2 = \lambda\gamma^1 \tag{A9}$$

$$c^1_1 : u'(c^1_1)\left(\gamma^1 + \mu_3 + \mu_4 - \mu_1\right) = \lambda\gamma^1 \tag{A10}$$

$$y^2_1 : v'(y^2_1/\theta^2)/\theta^2\left(\gamma^2 + \mu_1\right) = \lambda\gamma^2 \tag{A11}$$

$$c^2_1 : u'(c^2_1)\left(\gamma^2 + \mu_1\right) = \lambda\gamma^2. \tag{A12}$$

Combining (A1) and (A2), and combining (A3) and (A4), give

$$\left(1 - \frac{v'(y^1_1/\theta^1)}{u'(c^1_1)}\frac{1}{\theta^1}\right) = \mu_2 u'(c^1_1)\left(\frac{v'(y^1_1/\theta^1)}{u'(c^1_1)}\frac{1}{\theta^1} - \frac{v'(y^1_1/\theta^2)}{u'(c^1_1)}\frac{1}{\theta^2}\right). \tag{A13}$$

$$\left(1 - \frac{v'(y^2_1/\theta^2)}{u'(c^2_1)}\frac{1}{\theta^2}\right) = 0. \tag{A14}$$

Second period Recall the notation $F(\xi^1_{E-U}) = F^1_{E-U}$ (and similar for the other variables of this type). Define

$$\gamma^{0,1} = \gamma^0\left[\{\pi^1(F^1_{E-U} - F^1_{P-E}) + \{\pi^1 F^1_{P-E} + \pi^2 F^2_{P-E^2}\}\right]. \tag{A15}$$

Note also that we can define the following compensated derivative for type 1:

$$\frac{d^1_{comp}\,F}{dy^1_2} = \frac{d\,F}{dy^1_2} + \frac{v'(y^1_2/\theta^1)}{u'(c^1_2)}\frac{1}{\theta^1}\frac{d\,F}{dc^1_2.} \tag{A16}$$

The period 1 analogue to (A13) then becomes

$$
\begin{aligned}
&\lambda\left(\gamma^1+\gamma^{0,1}\right)\left(1-\frac{v'(y_2^1/\theta^1)}{u'(c_2^1)}\frac{1}{\theta^1}\right)=\mu_2 u'(c_2^1)\left(\frac{v'(y_2^1/\theta^1)}{u'(c_2^1)}\frac{1}{\theta^1}-\frac{v'(y_2^1/\theta^2)}{u'(c_2^1)}\frac{1}{\theta^2}\right)\\
&-\gamma^0\pi^1 H^{1'}(\gamma^0\pi^1[1-F_{E-U}^1])\frac{d_{comp}^1 F_{E-U}^1}{dy_2^1}-\gamma^0\pi^2 H^{2'}(\gamma^0\pi^2[1-F_{E-U}^2])\frac{d_{comp}^1 F_{E^1-U}^2}{dy_2^1}\\
&-\lambda\gamma^0\frac{d_{comp}^1\ \left[\pi^1(F_{E-U}^1-F_{P-E}^1)\right]}{dy_2^1}\times\left(-c_E+(y_2^1-c_2^1)\right)\\
&-\lambda\gamma^0\pi^2\frac{d_{comp}^1\ \left[F_{E-U}^2-F_{P-E}^2)\right]}{dy_2^1}\times\left(-c_E+(y_2^2-c_2^2)\right)\\
&-\lambda\gamma^0\frac{d_{comp}^1\ \left[\pi^1 F_{P-E}^1+\pi^2 F_{P-E}^2\right]}{dy_2^1}\times\left((\alpha\theta^1-w_p)h_p+(y_2^1-c_2^1)\right)\\
&+\lambda\gamma^0\frac{d_{comp}^1\ \left[\pi^1(1-F_{E-U}^1)+\pi^2(1-F_{E-U}^2)\right]}{dy_2^1}\times(b_1+b_2).
\end{aligned}
\tag{A17}
$$

Note that the compensated derivatives in (A17) are zero. Too see this, note that

$$
\begin{aligned}
\frac{d\xi_{E-U}^i}{y_2^1}&=\frac{-\beta v'(y_2^i/\theta^1)/\theta^1}{v(e^i)}, i=1,2 & \frac{d\xi_{E-U}^i}{c_2^1}&=\frac{\beta u'(c_2^i)}{v(e^i)}, i=1,2\\
\frac{d\xi_{P-E}^1}{y_2^1}&=0 & \frac{d\xi_{P-E}^1}{c_2^1}&=0\\
\frac{d\xi_{P-E}^2}{y_2^1}&=\frac{-\beta v'(y_2^1/\theta^1)/\theta^1}{v(h_P)-v(e^2)} & \frac{d\xi_{P-E}^2}{c_2^1}&=\frac{\beta u'(c_2^1)}{v(h_P)-v(e^2)}.
\end{aligned}
$$

Therefore, equation (A17) reduces to read

$$
\lambda\left(\gamma^1+\gamma^{0,1}\right)\left(1-\frac{v'(y_2^1/\theta^1)}{u'(c_2^1)}\frac{1}{\theta^1}\right)=\mu_2 u'(c_2^1)\left(\frac{v'(y_2^1/\theta^1)}{u'(c_2^1)}\frac{1}{\theta^1}-\frac{v'(y_2^1/\theta^2)}{u'(c_2^1)}\frac{1}{\theta^2}\right). \tag{A18}
$$

Let us now proceed with the second period labor supply for type 2 individuals. By using

$$
\gamma^{0,2}=\gamma^0\pi^2(F_{E-U}^2-F_{P-E}^2), \tag{A19}
$$

the second period analogue to (A14) can be written as

$$
\begin{aligned}
&\lambda\left(\gamma^2+\gamma^{0,2}\right)\left(1-\frac{v'(y_2^2/\theta^2)}{u'(c_2^2)}\frac{1}{\theta^2}\right)\\
&-\lambda\gamma^2\pi^2\frac{d_{comp}^2\ \left(F_{E-U}^2-F_{P-E}^2\right)}{dy_2^2}\times\left(-c_E+(y_2^2-c_2^2)\right)\\
&-\lambda\gamma^2\pi^2\frac{d_{comp}^2\ F_{P-E}^2}{dy_2^2}\times\left((\alpha\theta^1-w_p)h_p+(y_2^1-c_2^1)\right).
\end{aligned}
\tag{A20}
$$

The compensated derivatives are zero also in this case, since

$$\frac{d\xi^2_{P-E}}{y^2_2} = \frac{\beta v'(y^2_2/\theta^2)/\theta^2}{v(h_P) - v(e^2)} \qquad \frac{d\xi^2_{P-E}}{c^2_2} = \frac{-\beta u'(c^2_2)}{v(h_P) - v(e^2)}$$

$$\frac{d\xi^2_{E-U}}{y^2_2} = \frac{-\beta v'(y^2_2/\theta^2)/\theta^2}{v(e^2)} \qquad \frac{d\xi^2_{E-E}}{c^2_2} = \frac{\beta u'(c^2_2)}{v(e^2)}.$$

Therefore, (A20) reduces to read

$$\left(1 - \frac{v'(y^2_2/\theta^2)}{u'(c^2_2)}\frac{1}{\theta^2}\right) = 0. \tag{A21}$$

A.2 Propositions 2–4

The proofs of Propositions 2-4 follow directly from the social first-order conditions for w_p, h_p, c_E , b_1, and b_2, presented in appendix A.3 below. More precisely, the policy rules for the wage and hours of work in public employment in Proposition 2 follow by reorganizing the social first-order conditions for w_p and h_p, respectively; the policy rule for the education benefit in Proposition 3 follows by reorganizing the social first-order condition for c_E, and the policy rules for the unemployment benefit in Proposition 4 follows by reorganizing the social first-order conditions for b_1 and b_2.

A.3 Social first-order conditions for c_E, b_1, b_2, w_p, and h_p

FOC for c_E

$$\begin{aligned}
&\phi u'(c_E)\left(\pi^1[F^1_{E-U} - F^1_{P-E}] + \pi^2[F^2_{E-U} - F^2_{P-E^2}]\right)\\
&-\gamma^0\pi^1 H^{1'}(\gamma^0\pi^1[1 - F^1_{E-U}])\frac{dF^1_{E-U}}{dc_E} - \gamma^0\pi^2 H^{2'}(\gamma^0\pi^2[1 - F^2_{E-U}])\frac{dF^2_{E-U}}{dc_E}\\
&-\lambda\gamma^0\left(\pi^1[F^1_{E-U} - F^1_{P-E}] + \pi^2[F^2_{E-U} - F^2_{P-E}]\right)\\
&+\lambda\gamma^0\frac{d}{dc_E}\left(\pi^1[F^1_{E-U} - F^1_{P-E}]\right)\left(-c_E + (y^1_2 - c^1_2)\right)\\
&+\lambda\gamma^0\pi^2\left(\frac{d}{dc_E}\left(F^2_{E-U} - F^2_{P-E}\right)\left(-c_E + (y^2_2 - c^2_2)\right) - \left(F^2_{E-U} - F^2_{P-E}\right)\right)\\
&+\lambda\gamma^0\frac{d}{dc_E}\left(\pi^1 F^1_{P-E} + \pi^2 F^2_{P-E}\right)\left([\alpha\theta^1 - w_P]h_P + (y^1_2 - c^1_2)\right)\\
&-\lambda\gamma^0\frac{d}{dc_E}\left(\pi^1[1 - F^1_{E-U}] + \pi^2[1 - F^2_{E-U}]\right)(b_1 + b_2) - \mu_4 u'(c_E) = 0
\end{aligned} \tag{A22}$$

FOC for $b_i (i = 1, 2, R)$

$$\begin{aligned}
&\phi\beta^{i-1}u'(b_i)\left(\pi^1[1-F^1_{E-U}] + \pi^2[1-F^2_{E-U}]\right) \\
&-\gamma^0\pi^1 H^{1'}(\gamma^0\pi^1[1-F^1_{E-U}])\frac{dF^1_{E-U}}{db_i} - \gamma^0\pi^2 H^{2'}(\gamma^0\pi^2[1-F^2_{E-U}])\frac{dF^2_{E-U}}{db_i} \\
&-\lambda\gamma^0\left(\pi^1[1-F^1_{E-U}] + \pi^2[1-F^2_{E-U}]\right) \\
&+\lambda\gamma^0\frac{d}{db_i}\left(\pi^1 F^1_{E-U} + \pi^2 F^2_{P-E}\right)(-c_E + (y^1_2 - c^1_2)) \\
&-\lambda\gamma^0\frac{d}{db_i}\left(\pi^1[1-F^1_{E-U}] + \pi^2[1-F^2_{E-U}]\right)(b_1 + b_2) = 0 \qquad \text{(A23)}
\end{aligned}$$

FOC for w_P

$$\begin{aligned}
&\phi h_P u'(w_P h_P)\left(\pi^1 F^1_{P-E} + \pi^2 F^2_{P-E}\right) \\
&-\lambda\gamma^0\pi^1\frac{d}{dw_P}\left(F^1_{E-U} - F^1_{P-E}\right)(-c_E + (y^1_2 - c^1_2)) \\
&-\lambda\gamma^0\,\pi^2\frac{d}{dw_P}\left(F^2_{E-U} - F^2_{P-E}\right)(-c_E + (y^2_2 - c^2_2)) \\
&+\lambda\gamma^0\Bigg(\frac{d}{dw_P}\left(\pi^1 F^1_{P-E} + \pi^2 F^2_{P-E}\right)([\alpha\theta^1 - w_P]h_P + (y^1_2 - c^1_2)) \\
&-h_P\left(\pi^1 F^1_{P-E} + \pi^2 F^2_{P-E}\right)\Bigg) - \mu_3 u'(w_P h_P)h_P = 0 \qquad \text{(A24)}
\end{aligned}$$

FOC for h_P The FOC for h_P is a bit more involved, since the derivative of V_P w.r.t. h_P depends on ξ. It takes the form

$$\begin{aligned}
&\phi\left(\pi^1\int_0^{\xi^1_{P-E}}[w_P u'(w_P h_P) - \xi v'(h_p)]f^1(\xi)d\xi + \pi^2\int_0^{\xi^2_{P-E}}[w_p u'(w_P h_P) - \xi v'(h_P)]f^2(\xi)d\xi\right) \\
&-\lambda\gamma^0\pi^1\frac{d}{dh_P}\left(F^1_{E-U} - F^1_{P-E}\right)(-c_E + (y^1_2 - c^1_2)) \\
&-\lambda\gamma^0\,\pi^2\frac{d}{dh_P}\left(F^2_{E-U} - F^2_{P-E}\right)(-c_E + (y^2_2 - c^2_2)) \\
&+\lambda\gamma^0\Bigg(\frac{d}{dh_P}\left(\pi^1 F^1_{P-E} + \pi^2 F^2_{P-E}\right)([\alpha\theta^1 - w_P]h_P + (y^1_2 - c^1_2)) \\
&+(\alpha\theta^1 - w_P)\left(\pi^1 F^1_{P-E} + \pi^2 F^2_{P-E}\right)\Bigg) - \mu_3(u'(w_P h_P)w_P - v'_P(h_P)) = 0 \qquad \text{(A25)}
\end{aligned}$$