Dondena Working Papers

Carlo F. Dondena Centre for Research on Social Dynamics and Public Policy

Welfare State and Taxation Unit

Anti-discrimination Legislation and the Efficiency-Enhancing Role of Mandatory Parental Leave

Spencer Bastani, Tomer Blumkin, and Luca Micheletto

Working Paper No. 88
May 2016

Università Bocconi • The Dondena Centre
Via Guglielmo Röntgen 1, 20136 Milan, Italy
http://www.dondena.unibocconi.it

ISSN-2035-2034

Anti-discrimination Legislation and the Efficiency-Enhancing Role of Mandatory Parental Leave*

Spencer Bastani[†] Tomer Blumkin[‡] Luca Micheletto[§]

May 20, 2016

Abstract

We study a setting where anti-discrimination legislation gives rise to adverse selection in the labor market. Firms rely on nonlinear compensation contracts to screen workers who differ in their family/career orientation. This results in a labor market equilibrium where career-oriented workers are offered an inefficiently low duration of parental leave. In addition, family-oriented workers are offered lower wages as compared to their equally skilled career-oriented counterparts. We demonstrate the usefulness of mandatory parental leave rules in mitigating the distortion in the labor market and derive conditions under which a Pareto improvement is possible. We also characterize the optimal parental leave policy and highlight the possibility for parental leave legislation to eliminate the wage penalty of family-oriented workers by supporting pooling employment contracts.

Keywords: anti-discrimination, adverse selection, parental leave, efficiency

JEL classification: D82, H21, J31, J83

*We are grateful to Dan Anderberg, Nils Gottfries, Oskar Nordström-Skans, Dan-Olof Rooth, as well as participants at the NORFACE Welfare State Futures Conference, The Uppsala Center for Labor Studies (UCLS) Annual Members Meeting, Umeå University, the CESifo Employment and Social Protection Conference, and the Annual Meeting of the Israeli Economics Association for helpful comments on an earlier draft of the paper.

[†]Department of Economics and Statistics, Linnaeus University; Linnaeus University Centre for Labor Market and Discrimination Studies; Uppsala Center for Fiscal Studies; Uppsala Center for Labor Studies, Sweden; CESifo, Germany. E-mail: spencer.bastani@lnu.se.

[‡]Department of Economics, Ben Gurion University, Israel; CESifo, Germany; IZA. E-mail: tomerblu@bgu.ac.il

[§]Department of Law, University of Milan, and Dondena Centre for Research on Social Dynamics and Public Policy, Bocconi University, Italy; Uppsala Center for Fiscal Studies, Sweden; CESifo, Germany. E-mail: luca.micheletto@unibocconi.it

1 Introduction

There is a growing empirical literature documenting wage penalties associated with parenthood. Workers who take a large share of responsibility for the caring of children tend to have less job experience, greater career discontinuity and shorter work hours, resulting in worse labor market outcomes as compared to workers who take a smaller share of this responsibility. In most countries it is primarily women who are absent from work for reasons relating to the care of their children. However, even though there are much fewer men who take substantial amounts of parental leave, the fathers who do take parental leave are likely to suffer even greater penalties than women given pre-existing gender-norms regarding the division of child care.[1]

Since women are responsible for the lions' share of the time spent on child care by parents, the empirical literature has mostly focused on the labor market penalties associated with motherhood. For women in the US, each additional child is associated, on average, with a wage penalty of around 5%. Interestingly, these penalties persist even after controlling for workplace factors and education (Waldfogel 1997, Budig and England 2001). Moreover, motherhood is regarded as one of the most important factors explaining gender-differences in labor market outcomes. Bertrand et al. (2010) followed Chicago MBA graduates during the years after graduation and analyzed the dynamics of gender-differences in earnings. They find that male and female MBAs have nearly identical labor incomes at the outset of their careers but then diverge the years following graduation due to differences in career interruptions and growing gender differences in weekly hours worked. While their study focuses on workers in the corporate and financial sector, they also present suggestive evidence using data from the Harvard and Beyond (H&B) project showing that female MBAs appear to have a more difficult time combining career and family than do, for example, female physicians.[2] The importance of the relationship between work flexibility and compensation has also been stressed by Goldin (2014), who finds that work flexibility is particularly costly for employers in the top of the job distribution.

In a recent paper, Stantcheva (2014) recognizes the importance of hard work as a way for employees to favorably influence the perceptions of their employers and thereby be eligible for a higher compensation. Stantcheva considers a setting where firms do not observe the productivity of workers and thereby have to rely on screening through nonlinear compensation contracts. While Stantcheva focuses on the de-

[1]For example, Albrecht et al. (2003) and Albrecht et al. (2015) find evidence that the negative effect of total parental leave on earnings in Sweden is even stronger for fathers than for mothers.

[2]Further evidence on the relationship between child-related absences and labor market outcomes is presented by Angelov et al. (2016). They find that 15 years after the first child has been born, the male-female gender gap increases by 10 percentage points, an effect they attribute to mothers' career interruptions in direct proximity to childbirth and to their long-term continuing responsibilities for child rearing e.g., by working part-time.

sign of optimal redistributive taxation, she also mentions the potential interaction between government regulatory policies and adverse selection/screening by firms. In particular, she notes that the nature of anti-discriminatory policies will impact the degree of adverse selection in the labor market and uses motherhood to exemplify her point. Stantcheva writes, "If direct discrimination against [women with children] is prevented as is the case in many countries firms will have to indirectly screen through the labour contract. They might then offer a menu of contracts: a low-paying, part-time contract with shorter hours and more maternity leave, likely to be taken up by working mothers, and a high-paying, full-time contract with overtime bonuses, late-afternoon and week-end meetings, and little parental leave, likely to be taken up by workers without small children." (p. 1319).

In this paper we present a theoretical model that captures the nonlinear relationship between compensation and flexibility that seems to be prevalent in the labor market. Our interest does not lie in the relationship between workplace flexibility and gender-equality, but rather in the wage penalties faced by both male and female workers who prefer flexible work contracts. Importantly, we are interested in the structure of labor contracts and the market inefficiencies that arise in the presence of anti-discrimination legislation that prevents firms from discriminating between workers based on variables (such as gender, age, or martial status) that would indicate workers' preferences for flexible working arrangements .

We focus on a particular aspect of workplace flexibility, namely, parental leave. For our purposes, parental leave will refer to the legal framework regulating the extent to which firms must grant their employees child-related absences from work. The most basic form of parental leave refers to the time parents are permitted to take off work in order to take care of a newborn child, but in many countries, parental leave extends beyond the care of infants, to encompass different aspects of workplace flexibility, such as allowing parents to take time off work to take care of an older child, or to take care of a sick child.[3]

[3]There are large differences across countries in terms of the generosity of parental leave. The United States is a country with one of the least generous systems. The vast majority of states in the US provide no paid parental leave at all and the extent to which labor contracts offer flexibility with respect to child-related absences is largely a decision made by employers. Parental leave in Europe, and especially in the Nordic countries is significantly more generous. According to the Parental Leave Directive of the European Union (2010/18/EU) parental leave allowances in EU countries must be at least four months for each parent. A country with one of the worlds' most generous systems is Sweden where each parent has the legal right to be absent from work until the child is 18 months old. In total, Swedish parents are entitled to 480 days of paid parental leave. In case the family does not exhaust the full 480 days within the first 18 months of becoming a parent, any remaining days can be saved, and used for parental leave spells up until the child is 8 years old. There is also a special rule which allows parents to take time off work to take care of a sick child. In fact, parents have the right to take up to 120 days off work per year for each sick child under the age of 12 in the household (and in special cases age 16). Thus, parental leave in Sweden extends far beyond the care of infants. In addition, parents in Sweden have the right to work 75% out of the normal working hours until the child is 8 years old (in Sweden a full-time worker spends on average 40 hours per week on the job).

Our model captures the segmentation of the labor market into different "tracks" that differ in terms of the possibilities offered to combine work and family life. We envision firms offering (i) family-oriented jobs that offer greater flexibility with respect to child-related absences from work but a lower compensation, and, (ii) career-oriented jobs that demand longer work hours but offer a higher compensation.[4]

We consider the realistic case where firms are not allowed to offer distinct contracts to different types of workers due to anti-discrimination legislation, implying that all workers choose from the same set of contracts. In this case firms behave *as if* they were operating under asymmetric information allowing us to employ tools developed in the seminal paper by Rothschild and Stiglitz (1976). In order to support a separating equilibrium, firms engage in profit maximization subject to an incentive-compatibility constraint that ensures that workers self-select into jobs appropriate with their type, as reflected by workers' family/career-orientation.[5] We proceed to show that the resulting labor market equilibrium is inefficient. In order to separate between the family- and career-oriented workers, the latter are offered a duration of parental leave lower than the efficient level.

Our contribution consists of two parts. First, we demonstrate that a system of mandatory parental leave can mitigate the distortion in the labor market and deliver a Pareto improvement.[6] Second, we derive the optimal welfare maximizing policy and show that mandatory parental leave may serve to eliminate the parenthood penalty through the implementation of a pooling equilibrium where different types of workers are offered the same labor contract.[7]

The details of our model are as follows. Firms offer bi-dimensional employment contracts that differ in terms of renumeration and the generosity of parental leave. Workers differ in their career/family-orientation, captured by the variation in the likelihood of using parental leave, which may reflect heterogeneity in preferences and/or nurturing capacities. Workers who have a higher likelihood of using parental leave are considered less productive from the perspective of the firm due to their greater expected workplace absence. If firms could, based on observable characteristics (such as, for instance, gender, age, marital status, number of dependent children), identify

[4]This segmentation of the labor market is consistent with the ideas in Gibbons and Murphy (1992).

[5]In this paper we focus on a screening model, however one could derive similar conclusions in a signaling model of work commitment.

[6]In this paper we focus on a novel role for a mandatory parental leave rule to correct an inefficiency that arises due to the nature of information transmission in the labor market. There are of course many other different possible reasons why the government would like to enact mandatory parental leave rules. The government might choose to intervene to internalize externalities associated with fertility and demographic composition, or with extended parental time with children at home; or, the government might chose to intervene on equity grounds, as a means to promote re-distributive goals, notably, to support gender equality. See also Summers (1989) for a discussion of the normative justification for enacting mandatory benefits.

[7]The possibility for maternal leave to reduce wage differences between mothers and non-mothers has previously been emphasized by Waldfogel (1998).

those workers who have a higher likelihood of being absent from the firm, those workers would, in a perfectly competitive labor market, be offered a contract with a lower compensation. However, if firms are not allowed to offer different contracts to workers who differ in their career/family-orientation due to anti-discrimination legislation, a distortion arises which is identical to the one due to adverse selection in models with asymmetric information. Thus, in the presence of anti-discrimination legislation firms have to offer one set of contracts that all workers are free to choose from, that is, they behave *as if* they were operating under asymmetric information, allowing us to use the Rothschild and Stiglitz (1976) equilibrium concept.

In this equilibrium, a market inefficiency arises as, in order to support a separating equilibrium, contracts offered to career-oriented workers must be distorted. In order to separate between career-oriented and family-oriented workers (who have different expected productivity from the perspective of the firm), career-oriented workers will be offered labor contracts with a high compensation but an inefficiently low amount of parental leave. Our central contribution is to show that enacting a mandatory parental leave rule, which dictates a minimum level of parental leave that all labor contracts must comply with, may increase labor market efficiency. A mandatory parental leave rule allows to mitigate the distortion in the market equilibrium by increasing the parental leave (and the utility) of career-oriented workers without affecting the parental leave generosity associated with contracts offered to family-oriented workers; at the same time, it enables to compensate the family-oriented workers for the resulting information rent that arises when contracts intended for career-oriented workers are made more generous with respect to parental leave (thereby maintaining incentive-compatibility).

We provide a characterization of the conditions under which a parental leave reform leads to a Pareto improvement and argue that recent trends in fertility rates and labor market participation strengthen the case for government intervention on efficiency grounds. We also discuss the generality of our findings and in particular the role of paid vs unpaid parental leave.

In addition to characterizing the efficiency-enhancing role of introducing a mandatory parental leave rule, we also analyze the socially optimal level of parental leave that maximizes a weighted average of the utilities derived by career- and family-oriented workers. We demonstrate that the optimal duration of parental leave increases with respect to the weight assigned to family-oriented workers in the social welfare function. Furthermore, we show that, when this weight is high enough, the social optimum is given by a pooling contract where all workers are offered the same level of compensation (and the same duration of parental leave). This implies that the parenthood penalty is fully eliminated.

The paper is organized as follows. In section 2 we outline our model and present the efficient laissez-faire allocation, where firms are allowed to discriminate in the labor market. In that section, we also present the anti-discrimination case which gives rise to adverse selection and which we use as our benchmark for our subsequent analysis. In section 3 we show how the government can achieve a Pareto-improvement by implementing a mandatory parental leave rule. Section 4 presents some comparative statics results, discusses existence of the labor market equilibrium, and presents a numerical example. This section also discusses the issue of paid parental leave and the connection to nonlinear income taxation. Section 5 characterizes the socially optimal parental leave policy, allowing for arbitrary welfare weights on the different types of workers. In that section, we also discuss the optimality of separating versus pooling equilibria from the perspective of social welfare maximization. Finally, section 6 offers concluding remarks.

2 Model

We consider a simple labor market with two types of workers, indexed by $i = 1, 2$, whose respective measures are given by $0 < \gamma^i < 1, i = 1, 2$. The total population mass is normalized to unity. Individuals are equally skilled in the labor market but are assumed to differ with respect to their likelihood of taking up parental leave which we denote by π^i where we assume $\pi^2 > \pi^1 > 0$. By focusing on agents that are equally skilled, we focus on the adverse selection problem that occurs in a particular segment of the labor market as firms attempt to screen equally skilled workers who differ in their career/family-orientation through the use of nonlinear compensation schemes.

The differences in π can either be attributed to variation in preferences, or reflect differences in ability to rear children/nurturing capacity (see Cigno 2011).

The utility function of a type i-worker is given by:

$$U^i(c^i, \alpha^i) = c^i + \pi^i v(\alpha^i), \tag{1}$$

where c denotes consumption and α denotes the duration of parental leave associated with having a child. The function v is assumed to be strictly increasing and strictly concave. The term $\pi^i v(\alpha^i)$ is the expected utility derived from parental leave. Parental leave contributes positively to utility based on the notion that there is a leisure component in parental leave or simply that parents enjoy spending time with their children.[8]

The output per unit of time (the marginal product of labor) and the time endow-

[8]Notice that we make several simplifying assumptions. We assume that there is no labor-leisure choice in the standard sense. A worker who does not take parental leave will spend her entire time endowment working. This is without loss of generality. The quasi-linear specification is invoked for tractability. All our qualitative results remain robust to assuming instead a strictly concave utility from consumption. Moreover, all our qualitative results remain unscathed with endogenous fertility.

ment of each agent are each normalized to unity. We assume a perfectly competitive labor market implying that the market wage rate (per unit of time allocated to work) is given by unity, remunerating each worker according to her marginal product. However, as we explain below, workers will be differentially productive from the perspective of the firm as they differ in their probability of child-related absences from work.

We consider the following type of labor contract. Each firm offers a bundle (y, α) where y denotes total compensation and α reflects the generosity of parental leave associated with the labor contract. We think of a labor contract offering a higher α as being associated with a longer total duration of parental leave. An equivalent interpretation would be that the contract offers a greater flexibility with respect to child-related absences from work. Workers will choose between less demanding jobs that allow for more time with the family but a lower compensation and more demanding jobs that offer less family time but with a higher compensation.

The quantity $\pi\alpha$ is the expected duration of parental leave for a π-type worker. Thus, although workers produce the same output per unit of time spent at the firm, the higher π is, the lower is the expected output from the worker.[9]

The differences between workers are reflected in the labor market segmentation between less demanding ('part-time') and more demanding ('full-time') jobs. The former give more flexibility with respect to child-related absences accompanied by modest compensation, and are chosen by family-oriented workers (type 2), whereas the latter offer less flexibility but higher compensation, and are chosen by career-oriented workers (type 1). Even though we do not present a formal model of family decision-making, assuming that the primary earner is always career-oriented and has a fixed level of income, one may also interpret our model as focusing on the career/family trade-off faced by the secondary earner.

Free entry implies that firms may only choose contracts that yield zero profits. A firm offering a contract to a type-i worker must satisfy

$$y^i = 1 - \pi^i \alpha^i \tag{2}$$

where $\pi^i \alpha^i$ is the expected time worker i will be away from work.

Due to anti-discrimination legislation, firms cannot condition contracts on π or on observable variables that are correlated with π (such as age, marital status or the number of dependent children). This gives rise to a distortion which is identical to the one that arises due to adverse selection in the presence of asymmetric information. Before turning to the adverse-selection case, we briefly describe the efficient laissez-faire labor

[9] The formulation that we use is very tractable since it allows us to use the same parameter π to capture both the fundamental career-family trade-off manifested in the different orientation of agents towards parental leave and that individuals with different π will have different productivity from the perspective of the firm.

market equilibrium that arises in the absence of anti-discrimination legislation.

2.1 Laissez-faire efficient equilibrium

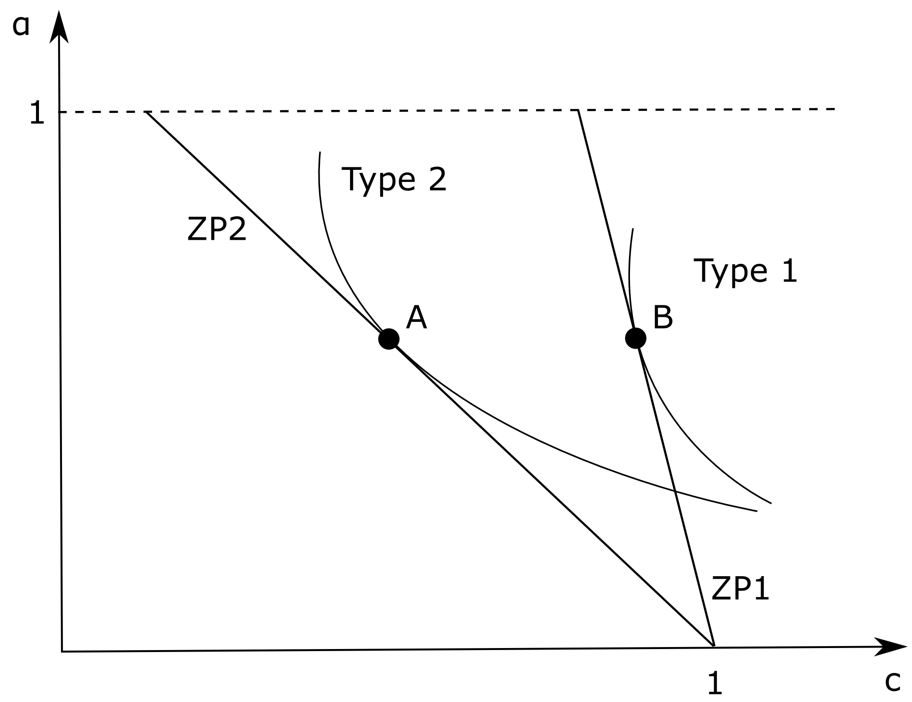
If firms were able to discriminate based on π, each worker would be offered a distinct contract that maximizes the utility in (1) subject to the budget constraint (2) resulting in an efficient labor market equilibrium.

The optimal contract for a type-i worker satisfies the familiar tangency condition given by:

$$\frac{1}{\pi^i v'\left(\alpha^i\right)} = \frac{1}{\pi^i} \quad \Longleftrightarrow \quad 1 = v'\left(\alpha^i\right) \tag{3}$$

The optimal contract is given by the solution to the system of two equations: the zero profit condition (budget constraint) in (2) and the MRS condition in (3). The optimum for type $i = 1,2$ is illustrated graphically in figure 1. Point A represents the contract offered to type-2 workers and point B represents the contract offered to type-1 workers. Note that because of the heterogeneity in π, agents have differently sloped budget- and indifference curves in the (c, α)-space.

Figure 1: Efficient equilibrium. Point A illustrates the efficient contract offered to type-2 workers and point B represents the efficient contract offered to type-1 workers.



Straightforward full differentiation of the system of equations given by (2) and (3) with respect to π, noting that $c = y$ in the absence of any taxes or tranfers, yields the following comparative statics: $c^1 > c^2, \alpha^1 = \alpha^2$ and $\pi^2\alpha^2 > \pi^1\alpha^1$.

In the next subsection we demonstrate that anti-discrimination legislation gives rise to adverse selection and an inefficient labor market equilibrium.

2.2 Equilibrium with anti-discrimination legislation

We turn now to analyze the case when firms are not allowed to offer separate contracts due to anti-discrimination legislation. As we will show below, the resulting equilibrium in the presence of anti-discrimination rules is similar to the equilibrium analyzed in the seminal paper by Rothschild and Stiglitz (RS) (1976) in the presence of asymmetric information. The crucial observation is that, in the presence of anti-discrimination legislation, firms behave *as if* they did not observe workers' types. From now on, we will refer to this as our benchmark equilibrium. Notice that we choose as our benchmark the equilibrium with anti-discrimination legislation rather than the efficient laissez faire allocation.

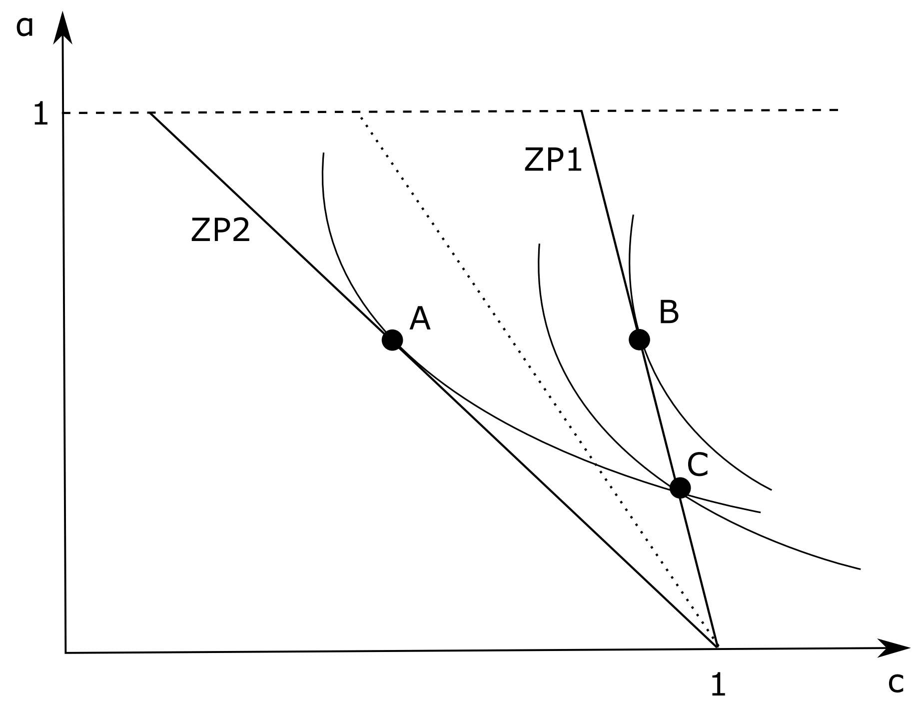
The *RS* equilibrium is defined by a set of labor contracts satisfying two properties: (i) firms make non-negative profits on each contract; and, (ii) there is no other potential contract that would yield non-negative profits if offered (in addition to the equilibrium set of contracts).

We focus on the separating equilibrium, which is illustrated in figure 2, along with the efficient equilibrium (where discrimination is allowed) described in the previous section. Notice that under the *RS* regime, as is well-known from Rothschild and Stiglitz (1976), a pooling equilibrium does not exist due to the potential for 'cream-skimming'. A separating equilibrium exists as long as the pooling line (i.e. the zero-profit line that would be relevant to firms hiring both types of workers), represented by the dashed line in figure 2, lies below the indifference curve of type-1 workers (as is the case in the figure). The issue of the existence of a separating equilibrium is discussed in the end of this section and further explored in section 4.

Notice that when the efficient contracts from section 2.1 (points A and B in the figure) are offered to both types of workers, both workers will prefer the contract intended for type-1 workers (point B in the figure). The pooling contract that would result when both workers pick the contract intended for type-1 would clearly yield negative profits to the firm (the point B lies above the zero profit line associated with pooling equilibrium allocations, given by $c = 1 - \alpha \sum \gamma^i \pi^i$). Hence, we conclude that this cannot be an equilibrium.

The separating equilibrium will maintain the efficient contract depicted by point A, which would still be offered to type-2 workers in the presence of anti-discrimination legislation. However, type-1 workers must be offered the contract depicted by point

Figure 2: Equilibrium in the presence of anti-discrimination legislation (benchmark). Type-2 workers are still offered their efficient contract A, whereas type-1 workers, due to the presence of the binding incentive constraint, must be offered contract C rather than the efficient contract B.



C in the figure, which lies on the intersection of the indifference curve of type-2 going though point A and the zero profit curve, associated with type-1 workers. Rather than maximizing the utility of type-1 worker subject to the zero profit condition (as happens in the efficient case), the new contract, C, maximizes the utility of type-1 subject to both the zero profit condition and the binding incentive constraint of type-2 workers, ensuring that type-2 workers would be indifferent between choosing point A and mimicking type-1 by choosing point C. The latter binding incentive constraint, that arises due to the presence of anti-discrimination legislation, is the source of inefficiency. Notice that the indifference curve of type-1 intersects (rather than being tangent to) the zero profit curve associated with type-1 workers. Thus, the resulting allocation implies that type-1 workers will work more hours, and correspondingly obtain a higher compensation, than under the laissez-faire equilibrium, yielding them a lower level of utility.[10]

For later purposes, we accompany the informal graphical illustration of this benchmark equilibrium with a formal definition:

Definition 1. *The labor market equilibrium in the presence of anti-discrimination legislation is given by the bundles* (c^{1*}, α^{1*}) *and* (c^{2*}, α^{2*}) *associated, correspondingly, with type 1 and type 2 workers, where* $c^{1*}, \alpha^{1*}, c^{2*}, \alpha^{2*}$ *solve the two zero profit conditions* $c^{i*} = 1 - \pi^i \alpha^{i*}, i = 1, 2,$ *the condition* $1 = v'(\alpha^{2*})$ *(the requirement that the bundle of type 2 is undistorted) and the condition* $c^{2*} + \pi^2 v(\alpha^{2*}) = c^{1*} + \pi^2 v(\alpha^{1*})$ *(the requirement that type 2 is indifferent between choosing her bundle and mimicking by choosing the bundle of type 1).*

Before turning to examine the potential efficiency enhancing role of government intervention we briefly discuss the issue of existence of a separating equilibrium and potential alternative equilibrium concepts.

2.2.1 Existence of a separating equilibrium

Recalling the definition of the RS equilibrium, one needs to rule out the possibility for a firm to offer a labor contract (in addition to the equilibrium set of contracts) that would yield *non-negative* profits. One possible scenario for a firm is to offer a separating contract that would be attractive for one type of workers only. However, this would be infeasible, as by construction, the separating equilibrium contracts maximize the utility of each type of worker subject to her respective binding budget constraint and (for type 1 workers) a binding incentive compatibility constraint associated with type-2 workers. Another possible scenario for a firm is to offer a pooling contract that would be attractive for both types of workers. As the indifference curve of type-1 worker is

[10] To see this formally, note that under full information, by virtue of condition (3), the allocation of type 1 workers satisfies $v'\left(\alpha^1\right) = 1$ whereas in the presence of asymmetric information the allocation of type-1 workers is distorted, implying that $v'\left(\alpha^1\right) > 1$. The result then follows by the strict concavity of v.

steeper than that of her type-2 counterpart at the separating type-1 bundle, a pooling allocation would be attractive for both types of workers if-and-only-if it would be attractive for type-1 workers. Thus, to rule out a profitable pooling offer, the zero-profit line associated with pooling equilibrium allocations (illustrated by the dashed line in figure 2) has to lie below the indifference curve of type-1 workers going through their separating equilibrium allocation. Formally, to ensure existence, we henceforth make the following assumption:

Assumption 1.

$$\max_{\alpha} 1 - \alpha \sum \gamma^i \pi^i + \pi^1 v(\alpha) < c^{1*} + \pi^1 v(\alpha^{1*}),$$

where (c^{1*}, α^{1*}) *denotes the type-1 bundle associated with the separating benchmark equilibrium.*

Assumption 1 implies that type-1 workers strictly prefer their separating equilibrium contract to any pooling contract that yields zero profits.

2.2.2 Alternative equilibrium concepts

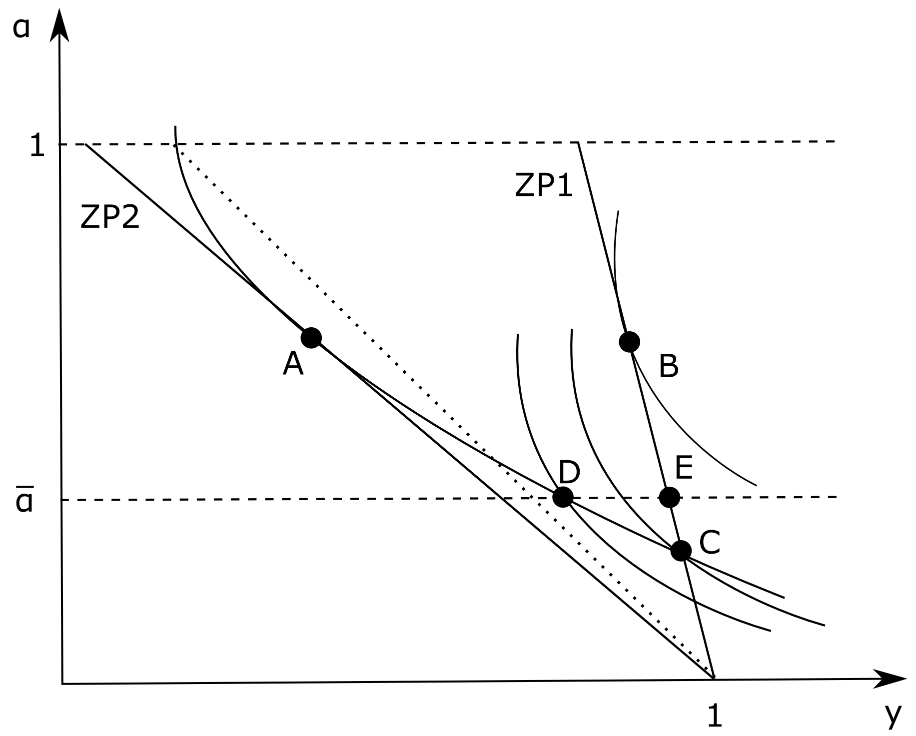
In this paper we focus on the RS equilibrium concept. A subsequent literature has challenged the negative prediction of RS, suggesting modified equilibrium concepts that may eliminate the market failure and hence give rise to second-best Pareto efficient allocations. One such notable example is the Miyazaki-Wilson-Spence (MWS) equilibrium [following Miyazaki (1977), Wilson (1977) and Spence (1978)]. The crucial difference between the two equilibrium concepts is in the degree of cross-subsidization across types that derives in equilibrium given the permissible forms of contracts that can be signed between the firms and the workers. Under the RS equilibrium concept each contract offered in equilibrium has to break even separately; under the alternative MWS equilibrium concept, instead, firms break even on their overall portfolio of contracts. Under the MWS concept, full cross subsidization is allowed and hence the resulting equilibrium is second-best Pareto efficient. As we wish to explore the role of government intervention in correcting the market failure associated with adverse selection, we need as our benchmark setting a framework which allows for less than full cross subsidization. For tractability we adopt the RS equilibrium concept.

3 Equilibrium with Parental Leave

The key question we wish to examine is whether the government can use its available policy tools to correct the market failure present in the benchmark equilibrium and thereby alleviate the adverse effects on labor market efficiency caused by anti-discrimination legislation.[11] We will focus on the potential efficiency-enhancing role played by a binding parental leave rule. Thus we assume the government sets a binding mandatory parental leave rule, denoted by $\bar{\alpha}$. That is, in equilibrium the following condition has to hold: $\alpha^i \geq \bar{\alpha}; i = 1, 2$.

The benchmark equilibrium analyzed in the previous section is illustrated as points A and C in figure 3. We recall two properties of the benchmark equilibrium: (i) the incentive constraint of type-2 agents is binding (in order to maintain incentive-compatibility type 1 workers have to be offered the point C rather than the efficient contract B) and (ii) the contract offered to type-2 agents is efficient. These two proper-

Figure 3: Equilibrium with parental leave. The contract depicted by point C in the figure is no longer feasible due to the presence of the parental leave rule.



ties of the benchmark equilibrium carry over to the equilibrium with parental leave.

The reason the incentive constraint of type-2 workers binds in the benchmark equilibrium is that, otherwise, firms could derive positive profits by offering contracts that would be attractive to type-1 workers only, by reducing work hours (increasing parental leave) and lowering the compensation. These types of profitable deviations

[11] Note that, due to the resulting adverse selection, the equilibrium allocation is clearly first-best inefficient. The question we turn to address is, however, whether this allocation is also second-best inefficient in light of anti-discrimination legislation and the policy tools available to the government.

are clearly not constrained by the presence of a parental leave rule.

The reason type-2 workers will obtain their efficient allocation is that, otherwise, firms can raise the utility of type-2 workers thereby creating a slack in the incentive-constraint. This would contradict property (i) above. As type-1 workers work longer hours than their type-2 counterparts in the benchmark equilibrium ($\alpha^2 > \alpha^1$), it follows that the parental leave rule will be slack for type-2 workers in equilibrium.

In figure 3 we have illustrated the introduction of a binding parental leave rule $\alpha = \overline{\alpha}$ that renders the point C infeasible (since it does not comply with the parental leave rule) but does not constrain the efficient contract offered to type 2 (point A). What is the equilibrium contract offered to type-1 workers in the presence of a binding parental leave rule? The fundamental difference between the benchmark allocation and the allocation arising in the presence of a parental leave rule is the following. In the benchmark regime, the allocation of type 1 worker is given by the intersection of the indifference curve of type 2 worker (going through her equilibrium allocation) and the zero profit line associated with firms hiring type-1 workers (point C in figure 3). In contrast, the allocation in a regime with a (binding) parental leave rule in place, is given by the intersection of the indifference curve of type 2 (going through her equilibrium allocation) and the parental leave rule line $\alpha = \bar{\alpha}$. This is illustrated by point D in figure 3.

Notice that since the parental leave rule is binding by assumption, the equilibrium contract offered to type 1 workers gives rise to positive profits for firms hiring type 1 workers. This is illustrated in figure 3 by virtue of the fact that point D lies below the zero profit line ZP1. The reason the contract offered to type-1 workers lies below their associated zero-profit line derives from the fact that the indifference curve associated with type-1 workers is steeper than that associated with their type-2 counterparts.

Notice that type 1 workers are made worse off when offered point D associated with the parental leave equilibrium as compared to being offered point C in the benchmark equilibrium. This is illustrated in the figure by the fact that the associated indifference curve going through point D lies below the indifference curve going through point C. However, since firms hiring type-1 workers derive positive profits in the presence of the parental leave rule, the government can tax these profits and rebate them back to agents in a lump-sum manner. If the size of the lump-sum grant is sufficiently large so as to bring the utility of type-1 agents to weakly exceed the benchmark level, a Pareto improvement is achieved (since type 2 agents would trivially be made strictly better off as compared to the benchmark equilibrium for any positive lump-sum transfer). To illustrate this graphically, notice that the lump-sum grant that is given to both types of workers implies an outward shift of the indifference curves of the two types of agents (going through points A and D). A Pareto improvement is achieved if the outward shift in the indifference curve of type 1 (going through point D) is sufficiently large so that the new indifference curve lies to the right of the indifference curve going through point C.

The possibility to obtain a Pareto improvement in the manner described above is proved formally in appendix A. Below we present a heuristic proof of this result using an intuitive argument. The idea is to start from the benchmark equilibrium, shifting the contract associated with type-1 workers along the zero profit line ZP1 in the direction of the efficient contract while compensating type-2 workers for the resulting information rent.

The argument proceeds as follows. Suppose we shift the contract offered to type 1 agent along the zero profit line ZP1 in the direction of the efficient contract (such as moving from point C to point E in figure 3). This shift would clearly make type-1 workers better off relative to the benchmark equilibrium. However, the point E would clearly not be incentive compatible. Type-2 workers would derive an information rent from such a shift since a more generous parental leave, reflected by a higher value of α, is valued more highly by type-2 workers who have a higher likelihood of using parental leave than their type-1 counterparts. This will lead to a violation of the type-2 agents' incentive constraint. Thus, to maintain the separating equilibrium incentive-compatible, type-2 workers need to be compensated for the resulting information rent. In order to keep the government's budget balanced, this compensation needs to be financed by some levy on type-1 workers. The government must, therefore, supplement the downwards shift in the work hours of type-1 workers with some form of cross subsidization from type-1 to type-2 workers. Clearly, this cross-subsidization increases the utility of type-2 workers beyond the benchmark level. To attain a Pareto-improvement, the utility of type-1 workers must therefore (weakly) exceed the benchmark level; namely, the (efficiency) gain from decreasing the work-hours of type-1 agents must outweigh the cost of compensating type-2 workers for the resulting information rent.

Let the profits associated with the contract offered to type-1 workers be denoted by $\sigma > 0$. Suppose that the government levies a confiscatory tax on the pure profits of firms hiring type-1 workers. Total tax revenues associated with this tax are given by $\gamma^1\sigma > 0$.

Assume further that these tax revenues are rebated back to agents in a lump-sum manner. As the population is normalized to unity, this (universal) lump-sum transfer is also equal to $\gamma^1\sigma$. Below we formally define the separating equilibrium associated with a parental leave rule, supplemented by pure profits taxation and a (universal) lump-sum transfer.

Definition 2. *The separating equilibrium associated with a parental leave rule, supplemented by pure profits taxation and a (universal) lump-sum transfer is given by the contracts* $(1 - \pi^1\overline{\alpha} - \sigma^*, \overline{\alpha})$ *and* $(y^2(\sigma^*), \alpha^2(\sigma^*))$ *where* σ^* *is the solution to:*

$$y^2(\sigma) + \gamma^1\sigma + \pi^2 v\left(\alpha^2\left(\sigma\right)\right) = 1 - \pi^1\overline{\alpha} - \sigma + \gamma^1\sigma + \pi^2 v\left(\overline{\alpha}\right), \tag{4}$$

and

$$\left\{y^2\left(\sigma\right), \alpha^2\left(\sigma\right)\right\} = \underset{y^2,\alpha^2}{\operatorname{argmax}}\, y^2 + \gamma^1\sigma + \pi^2 v\left(\alpha^2\right) \quad s.t. \quad y^2 = 1 - \pi^2\alpha^2 \tag{5}$$

In the above definition (5) states that type 2 workers receive their efficient contract along the zero-profit line $y^2 = 1 - \pi^2\alpha^2$, given the lump-sum transfer $\gamma^1\sigma$ whereas (4) states that the incentive constraint of type 2-workers is binding given the binding parental leave rule and the lump-sum transfer $\gamma^1\sigma$.

Notice that the net income on the right hand side of (4) is equal to the output produced by type-1 agents, namely $1 - \pi^1\overline{\alpha}$ (when restricted by the parental leave rule $\overline{\alpha}$), minus the pure profits σ plus the lump-sum transfer $\gamma^1\sigma$.

By virtue of the quasi-linear specification, $\alpha^2(\sigma) = \alpha^{2*}$, hence condition (4) simplifies to

$$1 - \pi^2\alpha^{2*} + \gamma^1\sigma + \pi^2 v(\alpha^{2*}) = 1 - \pi^1\overline{\alpha} - \gamma^2\sigma + \pi^2 v\left(\overline{\alpha}\right). \tag{6}$$

In addition to the simplified condition given in (6), to ensure the existence of an equilibrium associated with the parental leave rule, type-1 workers have to weakly prefer their separating equilibrium allocation to any pooling contract that yields zero profits. Formally, the following condition has to hold:

$$\max_{\alpha \geqslant \overline{\alpha}} 1 - \alpha \sum \gamma^i\pi^i + \gamma^1\sigma + \pi^1 v(\alpha) \leq 1 - \pi^1\overline{\alpha} - \gamma^2\sigma + \pi^1 v\left(\overline{\alpha}\right).$$

Notice that this condition is implied through continuity by assumption 1, provided that the degree of cross-subsidization induced by imposing the binding parental leave rule is sufficiently small.

It is straightforward to verify that by setting a binding parental leave rule, $\alpha^{1*} < \overline{\alpha} \leq \alpha^{2*}$, there exists a unique value of $\sigma > 0$ that solves condition (6). To see this first notice that when the parental leave rule is non-binding, namely $\overline{\alpha} = \alpha^{1*}$, then $\sigma = 0$, by construction of the benchmark equilibrium. Further notice that $\frac{\partial}{\partial\overline{\alpha}}[1 - \pi^1\overline{\alpha} + \pi^2 v\left(\overline{\alpha}\right)] > 0$, for all $\alpha^{1*} < \overline{\alpha} \leq \alpha^{2*}$, by virtue of the strict concavity of v and as $v'(\alpha^{2*}) = 1$ and $\pi^2 > \pi^1$. Thus, by setting a binding parental leave rule, namely $\alpha^{1*} < \overline{\alpha} \leq \alpha^{2*}$, the RHS of condition (6) will be larger than the LHS for $\sigma = 0$. Finally notice that by setting $\sigma = (\pi^2 - \pi^1)\overline{\alpha}/\gamma^2 > 0$ the LHS of condition (6) will be larger than the RHS, as

$1 - \pi^2\alpha^{2*} + \gamma^1\sigma + \pi^2 v(\alpha^{2*}) > 1 - \pi^2\overline{\alpha} + \pi^2 v(\overline{\alpha})$. Thus, by invoking the intermediate value theorem, continuity implies that there exists some $0 < \sigma < (\pi^2 - \pi^1)\overline{\alpha}/\gamma^2$ that solves condition (6). As the RHS is strictly decreasing in σ and the LHS is strictly increasing in σ, the solution is unique.

To sum up, imposing a binding parental leave rule, supplemented with pure profits taxation and a universal lump-sum transfer, provides exactly those features that are required to (potentially) achieve a Pareto improvement; namely, (i) a reduction in the work hours of type-1 workers which mitigates the distortion that arises due to anti-discrimination legislation, and, (ii) cross-subsidization between type-1 and type-2 workers that enables to compensate type-2 workers for the resulting information rent.

We turn next to characterize the necessary and sufficient conditions for such a composite policy reform to attain a Pareto improvement.

Proposition 1. *A Pareto improvement exists if-and-only-if*

$$\gamma^2/\gamma^1 < \frac{[v'(\alpha^{1*}) - 1]}{v'(\alpha^{1*})(\pi^2/\pi^1 - 1)},$$

where $\alpha^{i*}, i = 1, 2$, are associated with the separating benchmark equilibrium.

Proof See appendix A. □

The above proposition highlights that when the extent of induced cross-subsidization is small (γ^2 is small) and/or the adverse selection distortion is large (α^{1*} is small) the case for parental leave becomes stronger. The effect of differences in π on the above condition is generally ambiguous. We discuss this in detail in section 4.1.

The proof of the above proposition and all subsequent formal arguments are relegated to appendix A. Here we provide an intuitive informal derivation (heuristic proof) of the proposition using a perturbation argument.

We start out by noting that the contract offered to type-1 workers lies on the zero-profit line associated with firms hiring these workers. That is, the following condition is satisfied:

$$dy^1/d\alpha^1 = -\pi^1.$$

Moreover, at the benchmark separating equilibrium, agents of type 1 work more than the efficient amount of labor. This implies that their marginal willingness to pay for an increased α is larger than π^1:

$$MWP^1_\alpha = \pi^1 v'\left(\alpha^{1*}\right) > \pi^1.$$

Suppose that agents of type 1 are offered a compensated increase in α (compensated in the sense that their utility is kept unchanged via a proper reduction in consumption) and the firm gets π^1 in order to keep its zero-profit condition satisfied. Due to the distortion associated with the benchmark equilibrium allocation the government can collect from agents of type 1 an amount given by:

$$T^1 = \pi^1 \left[v' \left(\alpha^{1*} \right) - 1 \right] > 0.$$

Given that the proportion of agents of type 1 is γ^1, the revenue collected from agents of type 1 can then be used to finance a per-capita transfer to agents of type 2, which, assuming balanced budget, is given by:

$$T^2 = \frac{\gamma^1}{\gamma^2} T^1 = \frac{\gamma^1}{\gamma^2} \pi^1 \left[v' \left(\alpha^{1*} \right) - 1 \right].$$

For a mimicking type 2 agent, choosing the contract of type 1, utility is raised by:

$$\left(\pi^2 - \pi^1 \right) v' \left(\alpha^{1*} \right) > 0, \tag{7}$$

where the term measures the difference in the marginal willingness to pay for an increase in α between type 2 mimickers and a type 1 agents, and reflects an information rent. For a non-mimicking type 2 agent, choosing the contract associated with her type, utility is raised by:

$$T^2 = \frac{\gamma^1}{\gamma^2} T^1 = \frac{\gamma^1}{\gamma^2} \pi^1 \left[v' \left(\alpha^{1*} \right) - 1 \right] > 0. \tag{8}$$

Comparing (7) and (8), following some re-arrangements, it follows that mimicking by agents of type 2 will be discouraged when the following condition is satisfied:

$$\gamma^2 / \gamma^1 < \frac{\left[v' \left(\alpha^{1*} \right) - 1 \right]}{v' \left(\alpha^{1*} \right) \left(\pi^2 / \pi^1 - 1 \right)}. \tag{9}$$

The condition given in (9) replicates that stated in the proposition.

Notice that when condition (9) holds, the suggested policy reform, comprised of a compensated increase in α^1 supplemented by a transfer offered to type-2 workers that maintains the budget balanced, creates a slack in the incentive constraint associated with type-2 workers. The government can therefore reduce T^2 (and correspondingly adjust T^1 to maintain the budget balanced) up to the point where type-2 workers are just indifferent between choosing their own bundle and mimicking their type-1 counterparts. This shift would increase the utility of type-1 workers beyond the level associated with the benchmark equilibrium and would therefore give rise to a strict Pareto improvement (the utility of type-2 workers clearly increases due to the resulting information rent). The resulting allocation can be implemented by setting a mandatory binding parental leave rule, supplemented by confiscatory pure-profits taxation and a (universal) lump-sum transfer. This is shown formally in the appendix.

Further notice that condition (9) is both a necessary and sufficient condition for attaining a Pareto improvement which relies on the characteristics of the benchmark separating equilibrium. The right-hand side of (9) is independent of the ratio γ^2/γ^1 and defines an upper bound on the fraction of type-2 workers for a *Pareto* improvement to be feasible. The smaller is the fraction of type-2 workers (γ^2), the lower is the tax needed to maintain the incentive-compatibility constraint of type-2 workers while maintaining budget balance. This implies that an increase in the number of career-oriented workers relative to their family-oriented counterparts, i.e. a decrease in γ^2/γ^1, unambiguously makes a Pareto improvement more likely.[12]

In light of existing empirical evidence regarding the increased labor force participation of secondary earners and declining fertility rates, and to the extent that these trends are attributed to changing behavior among traditional (family-oriented) workers captured by a compositional change (a decrease in γ^2), then the case for government intervention on efficiency grounds becomes stronger.

A final remark regarding the necessity of condition (9) to achieve a Pareto improvement is in order. We have assumed the existence of a separating benchmark equilibrium and showed that the introduction of the parental leave system will necessarily make type 1 agents worse off in the new separating equilibrium with parental leave if condition (9) is not met. It is well known that in the RS 1976 setting, a pooling equilibrium does not exist. In the context of our model, a pooling benchmark equilibrium is not possible because if type 1 and type 2 workers were to be pooled at the same contract, a new firm could enter the market and offer a contract with slightly less α and a higher compensation, thereby attracting the more productive type 1 workers and derive positive profits. However, in the presence of a binding parental leave rule, such 'cream-skimming' by firms is not possible and a pooling equilibrium can be supported. This is in fact a novelty in our setting. However, switching from the benchmark equilibrium to a pooling equilibrium can never yield a Pareto improvement since by Assumption 1, *any* pooling equilibrium would necessarily make type-1 workers worse off compared to their benchmark allocation. Thus condition (9) is indeed both a necessary and sufficient condition to achieve a Pareto improvement.[13]

[12] Provided that this ratio does not fall below a certain threshold so that the separating equilibrium ceases to exist, see the discussion below and section 4.1.

[13] A pooling equilibrium supported by a parental leave rule can however be optimal from a social welfare perspective, as demonstrated in section 5.

4 Discussion and Extensions

In section 3 we saw that an increase in the number of career-oriented workers relative to their family-oriented counterparts, i.e. γ^2/γ^1, unambiguously makes a Pareto improvement more likely. In section 4.1 below we show that the extent to which changes in the differences in the π of the two types of agents make a Pareto improvement more or less likely is generally ambiguous. We also present a numerical example to resolve this ambiguity given certain parametric assumptions. The numerical example in section 4.1 also serves to demonstrate that it is possible to simultaneously satisfy the existence condition discussed in section 2.2.1 and the condition for Pareto improvement (9) for a wide range of parameter values.

In section 4.2 we also discuss the distinction between paid and unpaid parental leave, and in section 4.3 we explore the combination of nonlinear income taxation and mandatory parental leave.

4.1 Comparative statics with respect to π

We now examine the effects of changes in the differences in the likelihood of parental leave (the relationship between π^1 and π^2). For concreteness, we do this by fixing π^2 and consider changes in π^1.

Recall that condition (9) was expressed in terms of the quantities characterizing the market equilibrium with anti-discrimination legislation. Definition 1 states that in this equilibrium, the zero-profit conditions are satisfied, the bundle of type 2 is undistorted, and type 2 is indifferent between choosing her own contract and choosing the contract associated with type 1. Formally, this implies that $v'(\alpha^2) = 1$ and $c^2 + \pi^2 v(\alpha^2) = c^1 + \pi^2 v(\alpha^1)$. Insertion of the zero profit (budget) constraints (2), $1 - \alpha^2\pi^2 = c^2$ and $1 - \alpha^1\pi^1 = c^1$, into the two equations defining the benchmark equilibrium yields:

$$v'(\alpha^2) = 1, \tag{10}$$

$$1 - \alpha^2\pi^2 + \pi^2 v(\alpha^2) = 1 - \alpha^1\pi^1 + \pi^2 v(\alpha^1). \tag{11}$$

Now fix π^2 and consider (11). Since α^2 is given by the implicit solution to (10), the LHS of (11) expression does not depend on π^1. Total differentiation of (11) with respect to π^1 yields:

$$0 = \left[-\alpha_1 - \pi^1 \frac{\partial \alpha^1}{\partial \pi^1}\right] + \pi^2 v'(\alpha^1)\frac{\partial \alpha^1}{\partial \pi^1}.$$

This can be re-arranged as

$$\alpha_1 = \frac{\partial \alpha^1}{\partial \pi^1}\left[\pi^2 v'(\alpha^1) - \pi^1\right]. \tag{12}$$

The fact that $\pi^2 > \pi^1$ and that $v'(\alpha^1) > 1$ (stemming from the fact that the bundle of type 1 is distorted such that she works more than the efficient amount) implies that:

$$\frac{\partial \alpha^1}{\partial \pi^1} > 0 \quad \text{and} \quad \frac{\partial c^1}{\partial \pi^1} < 0. \tag{13}$$

Consider now expression (9). We can rewrite this expression as:

$$\gamma^2/\gamma^1 < \frac{\left[1 - \frac{1}{v'(\alpha^1)}\right]}{\frac{\pi^2}{\pi^1} - 1}. \tag{14}$$

It can immediately be seen that for π^2 fixed, a decrease in π^1 implies that the denominator in (14) increases which works in the direction of making it less likely for the government to achieve a Pareto improvement. Moreover, we know from (13) that a decrease in π^1 implies that α^1 decreases. Thus, the numerator $\left[1 - \frac{1}{v'(\alpha^1)}\right]$ in (14) increases by virtue of the strict concavity of v, which works in the direction of making it more likely for the government to attain a Pareto improvement. This means that the sign of the effect of a decrease in π^1 on (14) is generally ambiguous, and therefore one cannot determine whether a decrease in π^1 makes it more or less likely for the government to attain a Pareto improvement.

At first glance, the above ambiguity is surprising because one might expect that as the difference between π^1 and π^2 becomes larger, the distortion that arises due to anti-discrimination legislation increases and thus the scope for government intervention would be larger. This is captured by the effect of a decrease in π^1 on the numerator of (9).

However, even though a decrease in π^1 (conditional on holding π^2 fixed) implies that the distortion in the first best sense becomes larger, the information rent derived by type-2 workers becomes larger as well, as captured by the effect of a decrease in π^1 on the denominator in (9). The latter makes it more difficult for the government to intervene on efficiency grounds, rendering the total effect of a decrease in π^1 on expression (9) ambiguous.

To resolve this ambiguity we resort to a numerical example. This numerical example will also serve to illustrate the existence condition for a separating equilibrium discussed in section 2.2.1. As demonstrated by Rothschild and Stiglitz in their seminal paper, the separating equilibrium exists only when the fraction of type-2/type-1 workers in the population exceeds a certain threshold. This threshold ensures that type-1 workers strictly prefer the bundle associated with them in the benchmark separating market equilibrium to any bundle associated with a pooling allocation. We illustrate this lower bound in our numerical example. The details of the derivation of this lower bound can be found in appendix B.2.

For these purpose, we assume that the utility from parental leave is CRRA, $v(\alpha) = \frac{\alpha^b}{b}$, where $0 < b < 1$ to ensure concavity.

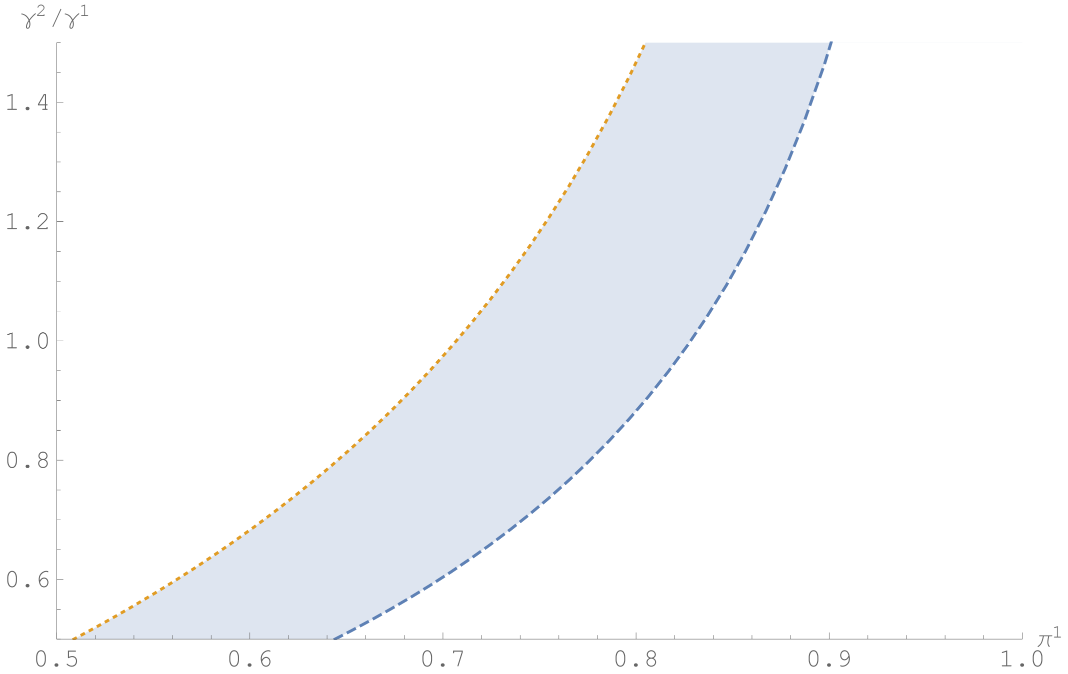


Figure 4: Numerical illustration of a region where the existence condition and the condition for Pareto-improvement are simultaneously satisfied.

In figure 4 we have plotted two upwards sloping curves. The lower curve represents the existence condition, which requires that for any π^1, the fraction of type-2 workers is sufficiently large to ensure existence of a separating equilibrium. The upper curve depicts condition (9) satisfied as an equality, which implies that for any π^1, a Pareto improvement is attainable if and only if the fraction of type 2 workers is sufficiently small. These curves separate the space into three distinct regions. The shaded region represents the set of parameter combinations for which a separating equilibrium exists and a Pareto improvement is attainable. In the lower region a separating equilibrium fails to exist, and in the upper region, the benchmark allocation is second best efficient. The figure demonstrates that a Pareto improvement is possible for a wide range of parameter combinations.[14]

[14]Notice that according to our parametric specification, the necessary and sufficient condition (14) for a Pareto improvement to exist, is homogeneous in the ratio π^1/π^2. Thus the fact that we fixed π^2 and conducted the comparative statics with respect to π^1 is of no substance for the qualitative results, provided that we satisfy the existence condition.

A close inspection of the figure reveals that, given our parametric assumptions, the information rent effect captured by the denominator of expression (9) dominates. This is reflected graphically by the fact that the upper boundary is increasing in π^1.[15] This implies that, as π^1 decreases, the government is less likely to attain a Pareto improvement. In the simulations we have chosen a value of b equal to 0.25. The qualitative results in the figure remain robust to the change in the degree of concavity of the function v measured by the constant coefficient of relative risk aversion, $1 - b$.

4.2 Paid parental leave

As mentioned in the introduction, in most OECD countries (US being the exception) mandatory parental leave is paid, namely the government is subsidizing the child-related absences from work mandated by law. In our setting we have assumed so far that mandatory parental leave was unpaid. In this section we turn to relax this assumption and examine the implications for the possibility to attain a Pareto improvement relative to the benchmark allocation with anti-discrimination legislation.

Fixing the duration of parental leave (per child), $\overline{\alpha}$, and denoting the (per-period) subsidy by $s > 0$, the paid parental leave is essentially equivalent to a child benefit equal to $s\overline{\alpha} \equiv b > 0$. Now suppose that the government is imposing a binding mandatory parental leave, $\overline{\alpha}$, and levies a confiscatory 100 percent tax on the pure profits derived by firms employing type-1 workers. Suppose further that the government is rebating the tax revenues back to the workers using a linear benefit scheme taking the form: $T = a + b\pi$, where $b > 0$ and π denotes the probability of taking a parental leave. Notice that the linear scheme implies that the level of benefit varies across the two types of workers. Further notice that a universal lump-sum transfer is captured by the special case where $b = 0$.

Consider the equilibrium associated with a parental leave rule, supplemented by 100 percent pure profits taxation and a linear benefit scheme of the form described above. Let $\sigma > 0$ denote the level of profits associated with an employer of type-1 workers under the parental leave regime. In equilibrium, the incentive compatibility constraint associated with type-2 workers must bind, namely: $1 - \pi^2\alpha^{2*} + \pi^2 v(\alpha^{2*}) + a + b\pi^2 = 1 - \pi^1\overline{\alpha} - \sigma + \pi^2 v(\overline{\alpha}) + a + b\pi^2$, where $b, \sigma > 0$ and α^{2*} is the efficient duration of parental leave associated with type-2 workers. Now suppose that the induced allocation associated with the parental leave rule yields a Pareto improvement relative to the benchmark benchmark regime. By virtue of the balanced budget condition (all tax revenues are rebated back to the workers via the linear bene-

[15] To see this, consider equation (9) satisfied as an equality. The upward slope of the upper curve in figure 4 implies that the RHS of condition (9) is increasing in π^1. As we already demonstrated that both the numerator and denominator of the RHS of (9) are decreasing in π^1, this implies that the effect associated with the denominator is prevailing.

fit scheme), it follows: $\gamma^1\sigma = \gamma^1(a + b\pi^1) + \gamma^2(a + b\pi^2)$. Re-arranging then yields: $a + b\pi^2 = \gamma^1\sigma + \gamma^1 b(\pi^2 - \pi^1) > 0$, where the inequality follows as $b > 0$, $\sigma > 0$ and $\pi^2 > \pi^1$. As type-2 workers obtain the efficient duration of parental leave and receive a positive transfer (as was just shown), their utility is strictly higher than that associated with the benchmark regime. Thus, for a Pareto improvement to hold it suffices that the following condition holds: $1 - \pi^1\overline{\alpha} - \sigma + \pi^1 v(\overline{\alpha}) + a + b\pi^1 \geq 1 - \pi^1\alpha^{1*} + \pi^1 v(\alpha^{1*})$. Namely, the utility derived by type-1 workers under the parental leave regime weakly exceeds the utility derived under the benchmark allocation.

Now, suppose that we replace the linear benefit scheme with a universal lump-sum transfer, maintaining the parental leave rule, $\overline{\alpha}$. Let $\sigma' > 0$ denote the level of profits associated with an employer of type-1 workers under the parental leave regime supplemented by a universal lump-sum transfer. Further let a' denote the universal lump-sum transfer. In equilibrium, the incentive compatibility constraint associated with type-2 workers must bind, namely: $1 - \pi^2\alpha^{2*} + \pi^2 v(\alpha^{2*}) + a' = 1 - \pi^1\overline{\alpha} - \sigma' + \pi^2 v(\overline{\alpha}) + a'$, where α^{2*} is the efficient duration of parental leave associated with type-2 workers. It is straightforward to verify that $\sigma' = \sigma$. By virtue of the balanced budget condition, $\gamma^1\sigma = a' > 0$. As type-2 workers obtain the efficient duration of parental leave and receive a positive transfer (as was just shown), their utility is strictly higher than that associated with the benchmark regime. To establish that the universal lump-sum transfer induces a Pareto improvement, recalling that the profits derived by employers of type-1 workers remain as under the linear benefit regime ($\sigma' = \sigma$), it suffices to show that $a' > a + b\pi^1$. By virtue of the balanced budget condition it follows: $\gamma^1\sigma = \gamma^1(a + b\pi^1) + \gamma^2(a + b\pi^2)$. Re-arranging then yields: $a + b\pi^1 = \gamma^1\sigma - \gamma^2 b(\pi^2 - \pi^1) < \gamma^1\sigma = a'$, where the inequality follows as $b > 0$ and $\pi^2 > \pi^1$. We conclude that any paid parental leave system that attains a Pareto improvement can be replaced by an unpaid parental leave system that also attains a Pareto improvement (for the same parameters). Thus, the option to provide a paid parental leave system does not expand the set of parameters for which a Pareto improvement (relative to the benchmark allocation) can be attained.

What seems to be somewhat surprising at a first glance is easily interpreted by noticing that in the benchmark equilibrium the incentive constraint associated with type-2 workers is binding. In order to expand the set of parameters for which a Pareto improvement is attained, one has to use policy tools that mitigate this incentive compatibility constraint, namely, rendering it less attractive for type-2 workers to mimic their type-1 counterparts. A paid parental leave system which is equivalent to a system of child benefits is more attractive for workers who are more likely to take a child-related absence from their jobs (namely, type-2 workers). Hence, such an arrangement is found more attractive by type-2 workers than by their type-1 counterparts, and therefore cannot serve to mitigate the former's binding incentive constraint. In appendix C we show that by allowing to tax children (rather than providing benefits) one can indeed expand the set of parameters for which a Pareto improvement can be attained.[16]

4.3 Nonlinear income taxation

Recall that a necessary condition for obtaining a Pareto improvement is to induce cross-subsidization from type-1 towards type-2 workers. One might envision that such cross-subsidization would be achievable using a nonlinear income tax. In this section we show that mandatory parental leave is in general desirable even in the presence of a nonlinear income tax. The simple intuition for this result is that a parental leave allows to better target the workers who are subject to distortions in the benchmark equilibrium.

Let $\left\{ \left(y^{1*}, \alpha^{1*}\right), \left(y^{2*}, \alpha^{2*}\right) \right\}$ be the set of contracts that are offered in the benchmark equilibrium, where $y^{j*} = 1 - \pi^j \alpha^{j*}$ (for $j = 1, 2$) denote the income paid by a firm to a worker choosing the contract associated with a parental leave spell of α^{j*}. Assume that condition (9) is satisfied, so that a binding parental leave rule, supplemented with pure profits taxation and a universal lump-sum transfer, can Pareto-improve upon the benchmark equilibrium. Denote respectively by $\overline{\alpha}$, with $\overline{\alpha} > \alpha^{1*}$, and $T > 0$ the length of the parental leave spell legislated by the government and the value of the uniform lump-sum transfer paid to all workers under a Pareto-improving public intervention scheme. At the new separating equilibrium the uniform lump-sum transfer paid by the government is financed by taxing the profits obtained by the firm employing type 1 workers. Thus, the post-intervention equilibrium set of labor contracts offered by firms will be given by: $\left\{ \left(y^{1*} - \left(\overline{\alpha} - \alpha^{1*}\right) \pi^1 - \frac{T}{\gamma^1}, \overline{\alpha}\right), \left(y^{2*}, \alpha^{2*}\right) \right\}$. Moreover, taking into account the uniform lump-sum transfer that everybody receives, the net-of-transfer consumption for type 1 workers will be $y^{1*} - \left(\overline{\alpha} - \alpha^{1*}\right) \pi^1 - \frac{T}{\gamma^1} + T = y^{1*} - \left(\overline{\alpha} - \alpha^{1*}\right) \pi^1 - \frac{\gamma^2}{\gamma^1} T$, and for type 2 workers it will be $y^{2*} + T$. Clearly, if this is the outcome that the government wishes to implement, a nonlinear income tax can be designed in such a way to induce the same outcome without any need to tamper with parental leave regulation. For instance, the government could design a nonlinear income tax such that workers would have to pay a huge tax for any level of earned income that is different than either $I^1 = y^{1*} - \left(\overline{\alpha} - \alpha^{1*}\right) \pi^1$ or $I^2 = y^{2*}$. Then, for anyone earning I^1 the associated income tax payment would be $T\gamma^2/\gamma^1$, whereas for anyone earning I^2 the associated income tax payment would be $-T$, i.e. an income transfer. With such a nonlinear income tax in place, firms would be forced to offer the same set of labor contracts as under the Pareto-improving parental leave scheme considered

[16] The potentially welfare enhancing role of taxing children has previously been recognized by Cigno and Pettini (2002).

above. It is important to notice, however, that the fact that nonlinear income taxation can implement the Pareto-improving scheme is not a general property. Rather, it is an artifact of the two-type setting that we have used to convey our central message.

In what follows we demonstrate how the government can expand the set of Pareto improving allocations by supplementing a non-linear tax and transfer system with a binding parental leave rule.

To see this, suppose that in addition to the two types of workers (1 and 2) there is a non-zero measure ($\gamma^0 > 0$, where γ^0 is assumed to be small) of workers, referred to as type-0, who derive no utility from parental leave, whose time endowment is normalized to unity and whose output per unit of time, denoted by y, distributes with some CDF over the support $[y^{0*}, y^{1*}]$, where $y^{0*} = 1 - \pi^1\alpha^{2*}$ and $y^{1*} = 1 - \pi^1\alpha^{1*}$. All variables designated with a star refer to the values prevailing in the benchmark equilibrium.

We assume that firms can readily distinguish between type 1 and 2 workers and their (lower skilled) type-0 counterparts as well as amongst type-0 workers. In the benchmark equilibrium, therefore, agents of type 0 would be offered a labor contract with no parental leave: $(y, 0)$. As type-0 workers are of a different skill level than the equally skilled type-1 and type-2 workers, anti-discrimination legislation (and hence the incentive compatibility constraints) will only apply to the latter two types.

We turn now to show that, in this setting, using non-linear taxation only, the government cannot implement an allocation which Pareto improves relative to the benchmark equilibrium allocation. Suppose, by way of contradiction that there exists an allocation that Pareto dominates the benchmark allocation. First notice that in such an allocation, $\alpha^{1*} < \alpha^1 \leq \alpha^{2*}$, namely the duration of parental leave of type-1 workers should be increased above the benchmark level to correct the distortion associated with the adverse selection. Further notice that to maintain the allocation incentive compatible type-2 workers have to be compensated for the resulting information rent. Denoting that tax levied on type-2 workers by T^2, it follows that in a Pareto dominating allocation $T^2 < 0$. Observe next that for any type-0 worker with income level $y \in [y^{0*}, y^{1*}]$, it must be the case that in a Pareto dominating allocation $T(y) \leq 0$ (otherwise the type-0 worker would be worse off relative to the benchmark allocation). In particular, consider the value $y^0 = 1 - \pi^1\alpha^1 \in [y^{0*}, y^{1*}]$. Then it is necessarily the case that $T(y^0) \leq 0$. However, recalling that the income level associated with type-1 workers is given by $y^1 = 1 - \pi^1\alpha^1$, it follows that $T(y^1) \leq 0$. Thus, in a Pareto dominating allocation none of the workers is paying positive taxes, where type-2 workers receive strictly positive transfers. It follows that the government runs into a deficit. We thus obtain the desired contradiction.

We have thus shown above that a nonlinear income tax cannot implement the Pareto-improving allocation in the extended setting. However, a Pareto-improvement can still be achieved by relying on a mandatory parental leave rule supplemented with pure profits taxation and a universal lump-sum transfer. To see this, notice that with a parental leave system one can avoid lowering the utility of type 0 agents since type 2 agents would only be cross-subsidized by type 1 agents. The revenue needed by the government to finance the uniform lump-sum transfer would only be collected from taxing the profits made by firms employing type 1 workers. Type 0 agents would in this case be made strictly better-off due to the fact that, as all other agents in the economy, they receive the uniform lump-sum transfer paid by the government.

Notice that the fact the type-0 workers receive a transfer implies that the lump-sum transfer is lower than in the case where the only agents in the economy are workers of type 1 and type 2. However, as, by virtue of condition (9), a strict Pareto improvement can be achieved in the two type case by setting the duration of the mandatory parental leave rule sufficiently close to α^{1*} (see the proof of proposition 1 in the Appendix), a Pareto improvement is obtained for the extended case, by continuity considerations, provided that the measure ot type-0 workers is sufficiently small.

Notice also that, when $\overline{\alpha}$ is introduced as a minimum parental leave spell that has to be part of all labor contracts offered in the economy, the contract offered by firms to type 0 agents will be $(y, \overline{\alpha})$. This is however of no harm for firms hiring type 0 workers since by assumption these workers do not value parental leave and will therefore not make use of this provision of the labor contract.

5 Welfare Maximization

In section 3 we have characterized a necessary and sufficient condition for a mandatory parental leave rule (supplemented by pure profits taxation and a universal lump-sum transfer) to be Pareto-improving relative to the benchmark allocation. In this section we turn to address the following normative question: what would be the socially desirable duration of parental leave? To answer this question we assume that social welfare is given by a weighted average of the utilities derived by both types of workers.

Our points of reference in this section are the durations of parental leave for the two types of agents in the benchmark allocation, α^{1*} and α^{2*} where $\alpha^{1*} < \alpha^{2*}$. By virtue of our previous analysis, we know that if condition (9) is satisfied, *marginally* introducing the parental leave system increases the utility of both types of agents. In this section we consider the welfare effects of introducing a *non-marginal* parental leave rule $\overline{\alpha} > \alpha^{1*}$. That is, we consider the effects of a parental leave rule that is binding for type 1 agents but may or may not be binding for type 2 agents.

To analyze the optimal duration of parental leave one must acknowledge that, depending on the value of $\overline{\alpha}$ set by the government, the government might be implementing either a separating or pooling labor market equilibrium. The possibility to attain a pooling equilibrium is in fact a novelty in our setting, as it is well-known, that in the standard RS 1976 setting, a pooling equilibrium does not exist. However, as explained in the end of section 3, the presence of a binding parental leave rule prevents 'cream-skimming' by firms and thereby may support a pooling equilibrium. Thus, to find the optimal parental leave policy we need to compare the social welfare levels for all types of labor market equilibria that can be supported. Thus, formally, social welfare is defined as follows:

$$W = \max_{j \in \{S,P\}} \left\{ \beta U_j^1(\overline{\alpha}) + (1-\beta) U_j^2(\overline{\alpha}) \right\}$$

where $U_j^i(\overline{\alpha})$ denotes the utility derived by a type i worker under an equilibrium of type $j = S, P$ (where S denotes the separating, and P denotes the pooling equilibrium) when the duration of parental leave is set to $\overline{\alpha}$. The parameter β denotes the weight type-1 workers carry in the social objective function. We also assume that any profits that may arise are taxed away and rebated back to agents in a lump-sum manner, in line with with section 3. To ease but slightly abuse notation, we will drop the subscript j in our exposition below, as well as in all the proofs in the appendix, as it will always be obvious which equilibrium regime that is under consideration.

We begin by characterizing the optimal duration of parental leave associated with a separating equilibrium. We then characterize the optimal duration of parental leave associated with a pooling equilibrium. Finally, we provide a general characterization of optimal parental leave policy by comparing the social welfare level attained in the optimal separating equilibrium with the social welfare level attained in the optimal pooling equilibrium for each level of the welfare weight β. In all our characterizations we assume that the necessary and sufficient condition for a Pareto-improvement (9) is satisfied.[17]

The optimal duration of parental leave under a separating equilibrium is characterized by the following proposition.

Proposition 2 (Separating Equilibrium)**.**

(i) The optimal solution under the separating regime is given by an interior solution $\overline{\alpha} \in (\alpha^{1*}, \alpha^{2*})$ *for* $\gamma^1 < \beta \leq 1$ *and by a corner solution,* $\overline{\alpha} = \alpha^{2*}$*, for* $0 \leq \beta \leq \gamma^1$*.*

(ii) For $\overline{\alpha} \in [\alpha^{1*}, \alpha^{2*}]$*,* $U^1(\overline{\alpha})$ *is strictly concave and* $U^2(\overline{\alpha})$ *is strictly increasing.*

(iii) Within the range of an interior solution, the optimal duration of parental leave under a separating equilibrium increases when β *decreases.*

Proof See appendix D □

[17] This assumption is not necessary but is made for simplicity. We comment on how it affects the results in footnote 18.

The proposition highlights the fact that, as the weight β assigned to workers with career-orientation decreases (with a corresponding increase in the weight attached to family-oriented workers), the optimal duration of parental leave increases. An increased duration of parental leave induces enhanced cross subsidization from career-oriented workers towards their family-oriented counterparts. As evident from part (ii) in the proposition, an increase in $\overline{\alpha}$ in the interval $[\alpha^{1*}, \alpha^{2*}]$ always raises the utility of type-2 workers, and, due to the efficiency-enhancing property of the mandatory parental leave rule, also initially raises the utility of type-1 workers. However, given the concavity of the utility of type-1 workers, there is a point when increasing the utility of type-2 workers comes at the expense of type-1 workers. This trade-off implies the possibility for an interior solution, depending on the value of β. When β is sufficiently small, we get a corner solution, and full cross-subsidization in the form of a pooling allocation becomes optimal.[18]

Notice that we have, just as in section 3, confined attention to the case where tax revenues (from the pure profits taxation of firms employing type-1 workers) are rebated via a uniform lump-sum transfer. Allowing for paid parental leave (see our discussion in section 4.2) would further enhance the government capacity to re-distribute from type-1 to type-2 workers. We then anticipate that the government will increase the generosity of the paid parental leave system as the weight assigned to family-oriented workers increases (alongside extending the duration of the parental leave).[19]

The next proposition characterizes the pooling regime.

Proposition 3 (Pooling Equilibrium)**.** *The optimal parental leave $\overline{\alpha}$ under a pooling equilibrium satisfies $\overline{\alpha} > \alpha^{1*}$, increases as β decreases, reaching $\overline{\alpha} = \alpha^{2*}$ when $\beta = \gamma^1$, and satisfies $\overline{\alpha} > \alpha^{2*}$ when $0 \leq \beta < \gamma^1$.*

Proof See appendix E $\square$

The proposition states that in the pooling equilibrium, as was the case in the separating regime, it is desirable to set a binding parental leave rule ($\overline{\alpha} > \alpha^{1*}$). Moreover, as was also the case in the separating equilibrium, the optimal duration of parental leave is an increasing function of the weight assigned to type-2 (family oriented) workers. Notably, as with the separating regime, a binding parental leave rule is desirable even for the limiting case where a full weight is assigned to type 1 (career oriented) workers,

[18]As mentioned on page 28, in our derivations we have assumed that the necessary and sufficient condition for Pareto improvement is satisfied. Without this assumption the characterization in proposition 2 would be qualitatively similar, barring the fact that the utility of type 1 would be monotonically decreasing with respect to the parental leave duration and that for high enough β, the optimum would be non-intervention (not setting a binding parental leave rule).

[19]When full weight is assigned to career-oriented workers, there will be nothing to gain from a paid parental leave structure, though, and the optimal system will remain one in which a universal lump-sum transfer is paid to both types of workers.

as it serves to mitigate the adverse selection distortion associated with the benchmark allocation. The higher the weight assigned to type-2 workers the longer is the duration of the parental leave rule, as the latter serves to enhance the degree of cross subsidization from type-1 to type-2 workers. The proposition also highlights the fact that when the weight attached to family-oriented workers is large, it is optimal to induce a pooling equilibrium with a duration of parental leave $\overline{\alpha}$ beyond α^{2*}, that is, beyond the point where it is binding for *both* types of agents. This implies that type-2 workers are actually taking more parental leave than the efficient amount. Nonetheless, when a relatively high weight is placed on the well-being of type-2 workers, raising $\overline{\alpha}$ above α^{2*} increases social welfare, as the higher duration of parental leave is more highly valued by type-2 workers $(\pi^2 > \pi^1)$ and there is implicit cross-subsidization from type-1 to type-2 workers.

We turn next to compare between the two regimes with the following proposition that characterizes the social optimum as a function of the weight assigned to type-1 (career-oriented) workers, β.

Proposition 4 (Characterization of the Social Optimum)**.**

(i) The separating allocation with $\overline{\alpha} \in (\alpha^{1*}, \alpha^{2*})$ *is the social optimum for* $\gamma^1 < \beta \leq 1$.

(i) The pooling allocation with $\overline{\alpha} \geq \alpha^{2*}$ *is the social optimum for* $0 \leq \beta \leq \gamma^1$.

(iii) The optimal duration of parental leave, $\bar{\alpha}(\beta)$*, is decreasing with respect to* β.

Proof See appendix F. □

Parts (i) and (ii) of the proposition establish that the social optimum is given by a separating equilibrium when the weight attached to career-oriented workers is relatively high, and that the social optimum is given by a pooling equilibrium when the weight attached to career-oriented workers is relatively low. Furthermore, part (iii) of the proposition summarizes the insights established in propositions 2 and 3, that the optimal duration of parental leave is increasing with respect to the weight assigned to family oriented workers (type-2). This reflects the desire of the government to redistribute towards family-oriented households, so as to mitigate the 'parenthood penalty'. In fact, in a pooling allocation (attained either as a corner solution for the separating regime, or as an interior solution for the pooling regime) the parenthood penalty is fully eliminated.[20]

[20]We would like to make a remark on the issue of implementability. Notice that when β is sufficiently low, the social optimum is a pooling allocation with $\overline{\alpha} > \alpha^{2*}$. For such values of $\overline{\alpha}$, a separating equilibrium cannot exist. However, in the case with a high β, $\overline{\alpha} < \alpha^{2*}$ and both the separating and pooling allocation can co-exist. Therefore, in order to achieve full implementation of the separating allocation, one needs to ensure that a pooling allocation cannot form an equilibrium. One way to do this would be to impose a 100 percent confiscatory income tax on the income level associated with the pooling allocation.

Before closing this section, we would like to mention that the possibility to attain a pooling equilibrium is a novelty in our setting. For this reason, it is useful to briefly relate to the possibility for bunching that has been highlighted in the optimal tax literature, initiated by Mirrlees (1971).

In the Mirrleesian optimal tax setting, it is assumed that firms (unlike the government) observe workers' types. With two types of workers [as in Stiglitz (1982)] bunching is never optimal (in fact it is Pareto dominated by the laissez-faire allocation). In contrast, our analysis suggests that pooling is socially desirable when the weight assigned to type-2 workers is sufficiently high. The reason for the difference derives from the fact that in a pooling allocation both types receive the same compensation (net income) but differ in the expected working time due to the difference in the likelihood of taking up parental leave. This implies that there is cross subsidization from type-1 workers (whose expected working time is higher than the average) to their type-2 counterparts (whose expected working time is lower than the average). In contrast, in the standard Mirrleesian framework, bunching implies that both type of workers obtain the same gross-income/net income bundle. Thus, there is no redistribution of income between the two types. The reason for the difference stems from the presence of anti-discrimination legislation that, in our setting, induces firms to behave as if they are operating under asymmetric information, unable to distinguish between the two types of workers.[21]

6 Concluding remarks

The general message of our paper is to highlight the potential for a mandatory parental leave rule to mitigate the distortions that arise in the labor market due to anti-discrimination legislation. These distortions arise when firms screen workers who differ in their career/family-orientation through nonlinear compensation contracts. To distinguish themselves from their family-oriented counterparts, career-oriented workers need to work more than the efficient amount and take too little parental leave. We have recognized that in the presence of anti-discrimination legislation, firms behave *as if* they were operating under asymmetric information, allowing us to use the Rothschild and Stiglitz (1976) equilibrium concept. In the context of a simple model we have characterized a necessary and sufficient condition for parental leave to be efficiency-enhancing and argued that recent trends in fertility rates and labor market participation strengthen the case for government intervention on efficiency grounds.

Maximizing a weighted average of the utilities derived by career- and family-oriented workers, we have also analyzed the socially optimal level of parental leave, highlight-

[21] For a related discussion, see Bastani et al. (2015).

ing that it might be associated with either separating or pooling employment contracts. The fact that the government can implement a pooling equilibrium is a novelty in our setting and demonstrates the potential for labor market regulation pertaining to work flexibility to mitigate the wage penalty faced by family-oriented workers in the marketplace. In particular, we have emphasized that, in the context of our model, a mandatory parental leave rule may completely eliminate the penalty associated with parenthood.

The efficiency-enhancing property of mandatory parental leave depends on the degree of adverse selection in the labor market. One way to quantify the extent of such adverse selection would be to use exogenous changes in anti-discrimination legislation that affect the abilities of firms to screen between workers. We leave this interesting empirical exercise for future research.

References

Albrecht, J., Björklund, A., and Vroman, S. (2003). Is There a Glass Ceiling in Sweden? *Journal of Labor Economics*, 21(1):145–177.

Albrecht, J., Thoursie, P. S., and Vroman, S. (2015). Parental Leave and the Glass Ceiling in Sweden : Gender Convergence in the Labor Market. *Research in Labor Economics*, 41:89 – 114.

Angelov, N., Johansson, P., and Lindahl, E. (2016). Parenthood and the Gender Gap in Pay. *Journal of Labor Economics*, 34(3).

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.

Bertrand, M., Goldin, C., and Katz, L. F. (2010). Dynamics of the gender gap for young professionals in the financial and corporate sectors. *American Economic Journal: Applied Economics*, 2(3):228–255.

Budig, M. J. and England, P. (2001). The Wage Penalty for Motherhood. *American Sociological Review*, 66(2):2004–225.

Cigno, A. (2011). Agency in family policy: A survey. *CESifo Economic Studies*, 57(2):305–331.

Cigno, A. and Pettini, A. (2002). Taxing family size and subsidizing child-specific commodities? *Journal of Public Economics*, 84(1):75–90.

Gibbons, R. and Murphy, K. J. (1992). Optimal Incentive Contracts in the Presence of Career Concerns: Theory and Evidence. *Journal of Political Economy*, 100(3):468.

Goldin, C. (2014). A Grand Gender Convergence: Its Last Chapter. *American Economic Review*, 104(4):1091–1119.

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

Miyazaki, H. (1977). The Rat Race and Internal Labor Markets. *Bell Journal of Economics*, 8(2):394–418.

Rothschild, M. and Stiglitz, J. (1976). Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information. *The Quarterly Journal of Economics*, 90(4):629–649.

Spence, A. M. (1978). Product differentiation and performance in insurance markets. *Journal of Public Economics*, 10(3):427–447.

Stantcheva, S. (2014). Optimal Income Taxation with Adverse Selection in the Labor Market. *Review of Economic Studies*, 81:1296–1329.

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

Summers, L. H. (1989). Some Simple Economics of Mandated Benefits. *The American Economic Review Papers and Proceedings*, 79(2):177–183.

Waldfogel, J. (1997). The Effect of children on women's wages. *American Sociological Review*, 62:209–217.

Waldfogel, J. (1998). Understanding the "Family Gap" in Pay for Women with Children. *Journal of Economic Perspectives*, 12(1):137–156.

Wilson, C. (1977). A model of insurance markets with incomplete information. *Journal of Economic Theory*, 16(2):167–207.

A Proof of proposition 1

We start with some preliminary useful definitions. A separating equilibrium allocation associated with a parental leave rule $\overline{\alpha}$, $\alpha^{1*} \leq \overline{\alpha} \leq \alpha^{2*}$, supplemented by a confiscatory tax levied on pure profits and a universal lump sum transfer, T, is given by: $\{\alpha^i, y^i\}_{i=1,2}, T$ where:

(i) $y^i = 1 - \pi^i \alpha^i; i = 1, 2,$

(ii) $\alpha^1 = \overline{\alpha},$

(iii) $\alpha^2 = \alpha^{2*}$, where $v'(\alpha^{2*}) = 1,$

(iv) $y^2 + T + \pi^2 v\left(\alpha^2\right) = y^1 - \frac{\gamma^2}{\gamma^1} \cdot T + \pi^2 v\left(\alpha^1\right),$

(v) $y^1 - \frac{\gamma^2}{\gamma^1} \cdot T + \pi^1 v\left(\alpha^1\right) \geq \max_{\alpha \geq \overline{\alpha}} 1 - \left(\sum \gamma^i \pi^i\right) \alpha + T + \pi^1 v(\alpha).$

Properties (iii) and (iv) carry over from the benchmark equilibrium implying that type-2 workers provide their efficient amount of labor [property (iii)] and that the incentive compatibility constraint associated with type-2 workers is binding [property (iv)]. Property (v) ensures that firms cannot offer a profitable pooling allocation that would be attractive for both types of workers by requiring that type-1 workers would weakly prefer their separating allocation to any pooling allocation that abides by the binding parental leave rule.

Substituting for α^i and y^i, $i = 1, 2$, from conditions (i)-(iii) into (iv) and re-arranging, yields: $T(\overline{\alpha}) = \gamma^1 \left(\pi^2 \alpha^{2*} - \pi^1 \overline{\alpha} + \pi^2 [v(\overline{\alpha}) - v(\alpha^{2*})]\right)$. Let $\hat{U}^1(\overline{\alpha})$ denote the utility derived by type-1 workers in the separating equilibrium associated with the parental leave rule, $\overline{\alpha}$. Formally, $\hat{U}^1(\overline{\alpha}) = 1 - \pi^1 \overline{\alpha} - \frac{\gamma^2}{\gamma^1} \cdot T(\overline{\alpha}) + \pi^1 v\left(\overline{\alpha}\right)$.

Lemma 1. *A Pareto improvement exists if-and-only-if there exists some $\overline{\alpha} > \alpha^{1*}$ for which $\hat{U}^1(\overline{\alpha}) \geq \hat{U}^1(\alpha^{1*})$.*

Proof Notice that $T(\alpha^{1*}) = 0$ by construction of the benchmark equilibrium. Further notice that T is strictly increasing with respect to $\overline{\alpha}$, by virtue of the strict concavity of v and the fact that $\overline{\alpha} \leq \alpha^{2*}$, $v'(\alpha^{2*}) = 1$ and $\pi^2 > \pi^1$. Thus, $T(\overline{\alpha}) > 0$ for all $\overline{\alpha} > \alpha^{1*}$. As type-2 workers provide their efficient amount of labor under any separating equilibrium [$\alpha^2 = \alpha^{2*}$ for all $\overline{\alpha}$] it follows that the utility derived by type-2 workers in any separating equilibrium associated with a binding parental leave rule, $\overline{\alpha} > \alpha^{1*}$, strictly exceeds their utility level associated with the benchmark allocation, $\overline{\alpha} = \alpha^{1*}$. Thus, a necessary and sufficient condition for obtaining a Pareto improvement relative to the benchmark allocation is that the utility derived by type-1 workers with a binding parental leave rule would weakly exceed their benchmark level of utility. This completes the proof. □

Lemma 2. *A Pareto improvement exists if-and-only-if the following condition holds:*

$$v'(\alpha^{1*})\left(\pi^1 - \gamma^2\pi^2\right) - \gamma^1\pi^1 > 0.$$

Proof Differentiating $\hat{U}^1(\overline{\alpha})$ with respect to $\overline{\alpha}$, evaluating the derivative at $\overline{\alpha} = \alpha^{1*}$, yields: $\left.\frac{\partial \hat{U}^1(\overline{\alpha}))}{\partial \overline{\alpha}}\right|_{\overline{\alpha}=\alpha^{1*}} = v'(\alpha^{1*})\left(\pi^1 - \gamma^2\pi^2\right) - \gamma^1\pi^1$. We turn to prove the sufficiency part first. Assume then that $v'(\alpha^{1*})\left(\pi^1 - \gamma^2\pi^2\right) - \gamma^1\pi^1 > 0$. By invoking a first-order approximation it follows that $\hat{U}^1(\overline{\alpha}) > \hat{U}^1(\alpha^{1*})$ for $\overline{\alpha}$ sufficiently close to α^{1*}. Notice further that by continuity considerations, property (v) in the definition of the separating equilibrium follows by virtue of assumption 1 and the fact that $T(\overline{\alpha}) \to 0$ as $\overline{\alpha} \to \alpha^{1*}$. Thus, we have constructed a well-defined separating allocation associated with a binding parental leave rule that Pareto dominates the benchmark allocation by virtue of lemma 1.

We turn next to the necessity part. Suppose then that $v'(\alpha^{1*})\left(\pi^1 - \gamma^2\pi^2\right) - \gamma^1\pi^1 \leq 0$. There are two separate cases to consider.

Suppose first that $\pi^1 - \gamma^2\pi^2 \leq 0$. It follows that $v'(\overline{\alpha})\left(\pi^1 - \gamma^2\pi^2\right) - \gamma^1\pi^1 < 0$ for all $\overline{\alpha} > \alpha^{1*}$. Thus, $\hat{U}^1(\overline{\alpha}) < \hat{U}^1(\alpha^{1*})$ for all $\overline{\alpha} > \alpha^{1*}$, hence, the benchmark allocation is second-best efficient by virtue of lemma 1. Suppose next that $\pi^1 - \gamma^2\pi^2 > 0$. Then, by virtue of the strict concavity of v, $v'(\overline{\alpha})\left(\pi^1 - \gamma^2\pi^2\right) - \gamma^1\pi^1 < 0$ for all $\overline{\alpha} > \alpha^{1*}$. Thus, $\hat{U}^1(\overline{\alpha}) < \hat{U}^1(\alpha^{1*})$ for all $\overline{\alpha} > \alpha^{1*}$, hence, the benchmark allocation is second-best efficient by virtue of lemma 1. □

Re-arranging the necessary and sufficient condition stated in lemma 2 yields that:

$$v'(\alpha^{1*})\left(\pi^1 - \gamma^2\pi^2\right) - \gamma^1\pi^1 > 0 \Longleftrightarrow \gamma^2/\gamma^1 < \frac{\left[v'\left(\alpha^{1*}\right) - 1\right]}{v'\left(\alpha^{1*}\right)\left(\pi^2/\pi^1 - 1\right)}.$$

This completes the proof of proposition 1.

B Details on the numerical example

B.1 The condition determining the existence of a Pareto improving allocation

Under the utility specification $v(\alpha) = \frac{\alpha^b}{b}, b > 0$, the conditions defining the benchmark equilibrium take the form:

$$\alpha^{2^{(b-1)}} = 1 \Longleftrightarrow \alpha^2 = 1 \tag{15}$$

$$u^2 = 1 - \pi^1\alpha^1 + \frac{\pi^2 {\alpha^1}^b}{b}, \quad \text{where} \quad u^2 = 1 - \pi^2\alpha^2 + \frac{\pi^2 {\alpha^2}^b}{b} = \frac{(1-b)\,\pi^2}{b} + 1 \tag{16}$$

$$c^1 = 1 - \pi^1\alpha^1, \tag{17}$$

$$c^2 = 1 - \pi^2\alpha^2 = 1 - \pi^2. \tag{18}$$

Notice that condition (15) determines the efficient amount of parental leave offered to type-2 workers; condition (16) is the binding (IC2) constraint which renders type-2 workers indifferent between mimicking type-1 or sticking to their contract, and conditions (17) and (18) state the consumption levels associated with type-1 and type-2 workers, respectively, determined by the corresponding zero profit conditions.

From (15)-(18) it can be derived that α^{1*} is given by the (unique) implicit solution to:

$$\frac{\alpha^1 b}{\left({\alpha^1}^b - (1-b)\right)} = \pi^2/\pi^1. \tag{19}$$

Thus, the necessary and sufficient condition for a Pareto improvement given in proposition 1, takes the form:

$$\gamma^2/\gamma^1 < \frac{1 - (\alpha^{1*})^{b-1}}{\pi^2/\pi^1 - 1}.$$

B.2 The condition determining the existence of a separating equilibrium

The critical threshold is the population ratio γ^2/γ^1 (satisfying $\gamma^1 + \gamma^2 = 1$) that makes type-1 workers just indifferent between the separating and the pooling allocations. This happens exactly when the pooling line is tangent to the indifference curve of type 1 workers in the separating equilibrium (see the dashed line in figure 2). Thus, the critical threshold is given by the implicit solution to the following system of equations:

$$\gamma^1 + \gamma^2 = 1, \tag{20}$$

$$\frac{1}{\pi^1 \alpha^{(b-1)}} = 1/(\gamma^1 \pi^1 + \gamma^2 \pi^2), \tag{21}$$

$$1 - \left(\gamma^1 \pi^1 + \gamma^2 \pi^2\right) \alpha + \frac{\pi^1 \alpha^b}{b} = 1 - \pi^1 \alpha^1 + \frac{\pi^1 {\alpha^1}^b}{b}, \tag{22}$$

where α^1 is the α for type 1 which prevails in the separating equilibrium and is given by the solution to (19).

Denoting the solution to (20)-(22) by the triplet $(\hat{\gamma}^1, \hat{\gamma}^2, \hat{\alpha})$, a separating equilibrium exists if-and-only-if:

$$\gamma^2/\gamma^1 \geq \hat{\gamma}^2/\hat{\gamma}^1. \tag{23}$$

C The Efficiency Enhancing Role of Taxing Children

In this section we demonstrate that by extending the set of policy instruments available to the government to include a non-linear tax and transfer system with means-tested child benefits which allows for taxing children, namely, the tax liability conditional on income is increasing with respect to the probability of taking up parental leave, one can expand the set of parameters for which a Pareto improvement can be attained.

As is common in the optimal tax literature, following the self-selection approach, we describe the tax-and-transfer system as two bundles associated with type-1 and type-2 workers, respectively, that satisfy a resource constraint and two incentive compatibility constraints.

Let the gross income associated with type-i worker ($i = 1, 2$) be denoted by y^i and let the corresponding tax (transfer if negative) be denoted by t^i. Consider the following tax system:[22]

$$\begin{aligned} y^1 &= 1 - \pi^1 \alpha^1, \\ y^2 &= 1 - \pi^2 \alpha^2, \\ t^1(\pi) &= (\pi - \gamma^1 \pi^1) z, \\ t^2 &= -\gamma^1 \pi^1 z, \end{aligned}$$

where:

i) $\alpha^1 = \alpha^{1*} + \epsilon$, with $\epsilon > 0$ and small,

[22] Notice that the tax function depends on the reported number of children π which varies between types. It makes type 2 (π^2) pay more taxes than type 1 (π^1) when mimicking, as $\pi^2 > \pi^1$.

ii) $\alpha^2 = \alpha^{2*}$,

and z is implicitly given by:

iii) $y^2 - t^2 + \pi^2 v(\alpha^2) = y^1 - t^1(\pi^2) + \pi^2 v(\alpha^1)$.

Several remarks are in order. First notice that two properties of the benchmark allocation carry over to the allocation induced by the tax-and-transfer system: type-2 workers provide their efficient amount of labor [condition (ii)] and the incentive constraint associated with type-2 workers is binding [condition (iii)]. Further notice that by construction of the benchmark allocation $1 - \pi^2\alpha^{2*} + \pi^2 v(\alpha^{2*}) = 1 - \pi^1\alpha^{1*} + \pi^2 v(\alpha^{1*})$. Thus, by substituting from the tax functions $t^i(\cdot), i = 1, 2$, into the binding incentive constraint associated with type-2 workers [given in condition (iii)], invoking a first-order approximation, $v(\alpha^1) = v(\alpha^{1*}) + \epsilon v'(\alpha^{1*})$, one can solve explicitly for z to obtain: $z = \epsilon[\pi^2 v'(\alpha^{1*}) - \pi^1]/\pi^2 > 0$, where the inequality sign follows as $\pi^2 > \pi^1$ and $v'(\alpha^{1*}) > 1$.

Finally notice that $z \to 0$ as $\epsilon \to 0$. Thus, by continuity considerations, by virtue of assumption 1, the incentive constraint associated with type-1 workers is satisfied (as a strict inequality). Formally,

$$y^1 - t^1(\pi^1) + \pi^1 v(\alpha^1) > y^2 - t^2 + \pi^1 v\left(\frac{1 - y^2}{\sum \gamma^i \pi^i}\right).$$

Notice the non-standard form of the incentive constraint [see the elaborate discussion in Stantcheva (2014) and Bastani et al. (2015)]. A standard incentive constraint would require that type-1 workers could not gain through mimicking their type-2 counterparts by choosing the bundle (α^2, y^2). This condition is in fact implied by the single crossing property (following from the fact that $\pi^2 > \pi^1$) and the fact that the incentive constraint associated with type-2 workers is binding. Instead, the constraint states that type-1 workers strictly prefer their separating equilibrium allocation to a pooling contract associated with the income level y^2 (associated with type-2 workers in the separating equilibrium) that yields zero profits.

To sum up, the tax system characterized above is incentive compatible. Furthermore, it can be readily verified that $\gamma^1 t^1(\pi^1) + \gamma^2 t^2 = 0$. Thus, the tax system maintains the budget balanced.

We turn next to provide a necessary and sufficient condition for the suggested tax system to attain a Pareto improvement relative to the benchmark allocation. Recalling that type-2 workers provide their efficiency amount of labor (as under the benchmark regime), α^{2*}, and further recalling that $z > 0$, it follows that the utility derived by type-2 workers under the tax system strictly exceeds that obtained under the benchmark regime. Thus, a necessary and sufficient condition for a Pareto improvement to exist is that the utility derived by type-1 workers under the tax system weakly exceeds their utility level under the benchmark regime. Formally, denoting the utility levels associated with the tax system and the benchmark regimes, correspondingly, by U^1 and U^{1*}, invoking a second-order approximation, $v(\alpha^1) = v(\alpha^{1*}) + \epsilon v'(\alpha^{1*}) + \epsilon^2 v''(\alpha^{1*})/2$, substituting for b and re-arranging, a Pareto improvement exists if-and-only-if:

$$
\begin{aligned}
&U^1 - U^{1*} = \\
&\qquad \epsilon \left(\pi^1[v'(\alpha^{1*}) - 1] - \gamma^2\pi^1[\pi^2 v'(\alpha^{1*}) - \pi^1]/\pi^2\right) + \pi^1\epsilon^2 v''(\alpha^{1*})/2 \geq 0 \\
&\qquad\qquad \iff v'(\alpha^{1*}) > (\pi^2 - \gamma^2\pi^1)/(\gamma^1\pi^2),
\end{aligned}
$$

where the last equivalence follows by virtue of the concavity of v and the fact that $\epsilon > 0$ is small (implying that the second order terms in the Taylor expansion can be ignored whenever $v'(\alpha^{1*}) \neq (\pi^2 - \gamma^2\pi^1)/(\gamma^1\pi^2)$).

Recall that by virtue of lemma 2 in the proof of the proposition, imposing a binding parental leave rule supplemented by a confiscatory tax levied on pure profits and a universal lump-sum transfer attains a Pareto improvement if, and only if, $v'(\alpha^{1*})\left(\pi^1 - \gamma^2\pi^2\right) - \gamma^1\pi^1 > 0$. Thus, in order to show that the suggested tax system expands the set of parameters for which a Pareto improvement exists, it suffices to demonstrate that the following condition is satisfied:

$$
\gamma^1\pi^1/\left(\pi^1 - \gamma^2\pi^2\right) > (\pi^2 - \gamma^2\pi^1)/(\gamma^1\pi^2). \tag{24}
$$

To render the analysis meaningful we assume that $\pi^1 - \gamma^2\pi^2 > 0$. Re-arranging the inequality condition given in (24), yields:

$$
\begin{aligned}
&\gamma^1\pi^1/\left(\pi^1 - \gamma^2\pi^2\right) > (\pi^2 - \gamma^2\pi^1)/(\gamma^1\pi^2) \\
&\qquad \iff \gamma^2(\pi^{1^2} + \pi^{2^2}) - (1 - \gamma^{1^2} + \gamma^{2^2})\pi^1\pi^2 > 0 \iff \\
&\qquad\qquad \gamma^2(\pi^2 - \pi^1)^2 > 0,
\end{aligned}
$$

where the last equivalence follows as $\gamma^{1^2} = (1-\gamma^2)^2 = 1 - 2\gamma^2 + \gamma^{2^2}$, which implies that $(1 - \gamma^{1^2} + \gamma^{2^2}) = 2\gamma^2$. This completes the proof.

D Proof of proposition 2

Let $U^i(\overline{\alpha}), i = 1, 2$, denote the type-i workers' utility level associated with the parental leave rule, $\overline{\alpha}$. By virtue of the definition of the separating equilibrium allocation associated with the parental leave rule, $\overline{\alpha}$ (see the proof of the proposition 1 for details), it follows:

$$
\begin{aligned}
U^1(\overline{\alpha}) &= 1 - \pi^1\overline{\alpha} - \frac{\gamma^2}{\gamma^1} \cdot T(\overline{\alpha}) + \pi^1 v\left(\overline{\alpha}\right), \\
U^2(\overline{\alpha}) &= 1 - \pi^2\alpha^{2*} + T(\overline{\alpha}) + \pi^2 v\left(\alpha^{2*}\right),
\end{aligned}
$$

where $T(\overline{\alpha}) = \gamma^1\left(\pi^2\alpha^{2*} - \pi^1\overline{\alpha} + \pi^2[v(\overline{\alpha}) - v(\alpha^{2*})]\right)$ denotes the universal lump-sum transfer associated with the parental leave rule, $\overline{\alpha}$.

Before turning to formulate the government problem, it is be useful to derive some comparative statics properties of the utility functions, $U^i(\overline{\alpha}), i = 1, 2$. We turn first to the utility of type-1 workers. Assuming that the necessary and sufficient condition for a Pareto improvement is satisfied, it follows that

$$
\left.\frac{\partial U^1(\overline{\alpha}))}{\partial \overline{\alpha}}\right|_{\overline{\alpha}=\alpha^{1*}} = v'(\alpha^{1*})\left(\pi^1 - \gamma^2\pi^2\right) - \gamma^1\pi^1 > 0.
$$

Namely, starting at the laissez-faire allocation, imposing a binding parental leave rule implies an increase in the utility of type-1 workers. The latter property furthermore implies that $\pi^1 - \gamma^2\pi^2 > 0$.

By virtue of the fact that $v'(\alpha^{2*}) = 1$, it follows that

$$
\left.\frac{\partial U^1(\overline{\alpha}))}{\partial \overline{\alpha}}\right|_{\overline{\alpha}=\alpha^{2*}} = v'(\alpha^{2*})\left(\pi^1 - \gamma^2\pi^2\right) - \gamma^1\pi^1 = -\gamma^2(\pi^2 - \pi^1) < 0.
$$

By virtue of the strict concavity of v it follows hence that there exists a unique value of α, which we denote by $\widetilde{\alpha}$, which satisfies

$$
\left.\frac{\partial U^1(\overline{\alpha}))}{\partial \overline{\alpha}}\right|_{\overline{\alpha}=\widetilde{\alpha}} = v'(\widetilde{\alpha})\left(\pi^1 - \gamma^2\pi^2\right) - \gamma^1\pi^1 = 0,
$$

such that for all $\alpha^{1*} \leq \overline{\alpha} < \widetilde{\alpha}$, $\frac{\partial U^1(\overline{\alpha}))}{\partial \overline{\alpha}} > 0$, whereas, for all $\widetilde{\alpha} < \overline{\alpha} \leq \alpha^{2*}$, $\frac{\partial U^1(\overline{\alpha}))}{\partial \overline{\alpha}} < 0$. We conclude that the utility of type-1 workers is strictly concave in the range $[\alpha^{1*}, \alpha^{2*}]$ and attains its maximum at $\widetilde{\alpha}$.

Turning next to the utility of type-2 workers, it follows that

$$
\frac{\partial U^2(\overline{\alpha}))}{\partial \overline{\alpha}} = \gamma^1[\pi^2 v'(\overline{\alpha}) - \pi^1] > 0,
$$

for all $\alpha^{1*} \leq \overline{\alpha} \leq \alpha^{2*}$, by virtue of the strict concavity of v and as $v'(\alpha^{2*}) = 1$ and $\pi^2 > \pi^1$.

The government optimization problem is given by:

$$\max_{\overline{\alpha}} \sum \beta^i U^i(\overline{\alpha}),$$

where $\sum \beta^i = 1$ and $0 \leq \beta^i \leq 1$. Formulating the first order condition with respect to $\overline{\alpha}$ yields (where we simplify notation by letting $\beta^1 \equiv \beta$) :

$$\begin{aligned} H(\beta, \overline{\alpha}) &\equiv \beta \frac{\partial U^1(\overline{\alpha}))}{\partial \overline{\alpha}} + (1-\beta)\frac{\partial U^2(\overline{\alpha}))}{\partial \overline{\alpha}} \\ &= \beta[v'(\overline{\alpha})\left(\pi^1 - \gamma^2\pi^2\right) - \gamma^1\pi^1] + (1-\beta)\gamma^1[\pi^2 v'(\overline{\alpha}) - \pi^1] \geq 0 \\ &\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (= 0, \overline{\alpha} < \alpha^{2*}). \end{aligned}$$

It is straightforward to verify that in case a full weight is assigned to type-1 workers ($\beta = 1$) then the optimal solution is interior and given by $\overline{\alpha} = \widetilde{\alpha}$. Alternatively, when a full weight is assigned to type-2 workers ($\beta = 0$) then the optimum is given by a corner solution, $\overline{\alpha} = \alpha^{2*}$, and the induced allocation is a pooling equilibrium in which both the duration of parental leave and the compensation is identical for both types of workers. Notice that, by construction, the duration of the parental leave rule under a separating allocation is bounded from above by α^{2*}.

When the optimum is obtained as an interior solution, then by virtue of the first-order condition with respect to $\overline{\alpha}$, recalling that $\frac{\partial U^2(\overline{\alpha}))}{\partial \overline{\alpha}} > 0$, it follows that $\frac{\partial U^1(\overline{\alpha}))}{\partial \overline{\alpha}} \leq 0$. Thus, $\partial H/\partial \beta < 0$. Moreover, by virtue of the strict concavity of v and the fact that $\pi^1 - \gamma^2\pi^2 > 0$, it follows that $\partial H/\partial \overline{\alpha} < 0$. Thus, $\partial \overline{\alpha}/\partial \beta = -\frac{\partial H/\partial \beta}{\partial H/\partial \overline{\alpha}} < 0$. Hence, within the range of an interior solution, the optimal duration of parental leave is increasing with respect to the weight assigned to type-2 workers (decreasing with respect to β).

As $v'(\alpha^{2*}) = 1$ and $\pi^2 > \pi^1$, it is straightforward to verify that $H(1, \alpha^{2*}) < 0$, $H(0, \alpha^{2*}) > 0$. Thus, by continuity considerations, the intermediate value theorem implies that there exists some $0 < \beta < 1$, denoted by $\hat{\beta}$, for which $H(\hat{\beta}, \alpha^{2*}) = 0$. Furthermore, it can be verified that $\frac{\partial H(\beta, \alpha^{2*})}{\partial \beta} = \pi^1 - \pi^2 < 0$, hence, $\hat{\beta}$ is unique. Substituting for $v'(\alpha^{2*}) = 1$ into the first-order condition $H(\hat{\beta}, \alpha^{2*}) = 0$, one can explicitly solve for the cutoff weight, $\hat{\beta}$, to obtain $\hat{\beta} = \gamma^1$.

Notice finally that as $\partial H/\partial \overline{\alpha} < 0$, the second-order condition for the government optimization problem is satisfied, so the optimum is indeed characterized by the first-order condition formulated above.

E Proof of proposition 3

Let $U^i(\overline{\alpha}), i = 1, 2$, denote the type-i workers' utility level associated with the parental leave rule, $\overline{\alpha}$. By construction of the mandatory parental leave rule, $\bar{\alpha} \geq \alpha^{1*}$. Furthermore, $U^i(\overline{\alpha}) = [1 - \overline{\alpha}(\gamma^1\pi^1 + \gamma^2\pi^2)] + \pi^i v(\overline{\alpha})$, $i = 1, 2$. Notice that, in contrast to the separating equilibrium, under the pooling regime expected profits are zero. Thus, there are no tax revenues and the lump-sum transfer is accordingly set to zero. Nonetheless, there is cross-subsidization between the two types of workers, as both receive the same level of compensation, but differ in the expected working time, due to the difference in the likelihood of taking up parental leave.

The government optimization problem is given by:

$$\max_{\overline{\alpha}} \sum \beta^i U^i(\overline{\alpha}),$$

where $\sum \beta^i = 1$ and $0 \leq \beta^i \leq 1$. Formulating the first order condition with respect to $\overline{\alpha}$ yields (where we again simplify notation by letting $\beta^1 \equiv \beta$) :

$$\begin{aligned} F(\beta, \overline{\alpha}) &\equiv \beta \frac{\partial U^1(\overline{\alpha}))}{\partial \overline{\alpha}} + (1-\beta) \frac{\partial U^2(\overline{\alpha}))}{\partial \overline{\alpha}} \\ &= -(\gamma^1\pi^1 + \gamma^2\pi^2) + [\beta\pi^1 + (1-\beta)\pi^2] v'(\bar{\alpha}) \leq 0 \quad (= 0, \overline{\alpha} > \alpha^{1*}). \end{aligned}$$

We first turn to show that, assuming that the necessary and sufficient condition for a Pareto improvement is satisfied, the welfare optimum under a pooling regime is always given by an interior solution; namely, $\overline{\alpha} > \alpha^{1*}$. To see this, one can re-arrange the first order condition to establish that a corner solution arises when the following inequality holds:

$$v'(\alpha^{1*}) \leq \frac{(\gamma^1\pi^1 + \gamma^2\pi^2)}{[\beta\pi^1 + (1-\beta)\pi^2]}.$$

At the same time, by virtue of the necessary and sufficient condition for a Pareto improvement, it follows that:

$$v'(\alpha^{1*}) > \frac{\gamma^1\pi^1}{(\pi^1 - \gamma^2\pi^2)}.$$

To demonstrate that a corner solution cannot exist, it suffices to show that

$$\frac{\gamma^1\pi^1}{(\pi^1 - \gamma^2\pi^2)} \geq \frac{(\gamma^1\pi^1 + \gamma^2\pi^2)}{[\beta\pi^1 + (1-\beta)\pi^2]},$$

which holds if-and-only-if (following some algebraic manipulations),

$$\gamma^1\beta{\pi^1}^2 + (1-\beta)\gamma^1\pi^1\pi^2 \geq \gamma^1{\pi^1}^2 + {\gamma^2}^2\pi^1\pi^2 - {\gamma^2}^2{\pi^2}^2.$$

Notice that the left-hand side of the above inequality expression is decreasing with respect to β, as $\pi^2 > \pi^1$. Thus, it suffices to prove that the inequality holds for $\beta = 1$.

Substituting for $\beta = 1$ yields upon re-arrangement: $\gamma^{2^2}\pi^{2^2} \geq \gamma^{2^2}\pi^1\pi^2$, which holds as $\pi^2 > \pi^1$. This completes the proof.

We conclude that the pooling optimum is given by an interior solution for all values of β.

Finally, notice that for $\beta = \gamma^1$, as $v'(\alpha^{2*}) = 1$, the optimal duration of parental leave is given by $\overline{\alpha} = \alpha^{2*}$. Further notice that by virtue of the strict concavity of v and the fact that $\pi^2 > \pi^1$, it follows that $\partial F/\partial\overline{\alpha} < 0$ and $\partial F/\partial\beta < 0$. Thus, $\partial\overline{\alpha}/\partial\beta = -\frac{\partial F/\partial\beta}{\partial F/\partial\overline{\alpha}} < 0$. Hence, the optimal duration of parental leave is increasing with respect to the weight assigned to type-2 workers (decreasing with respect to β).

Notice that as $\partial F/\partial\overline{\alpha} < 0$, the second-order condition for the government optimization problem is satisfied, so the optimum is indeed characterized by the first-order condition formulated above.

F Proof of proposition 4

Part (ii) Let $W^{sep}(\beta, \overline{\alpha})$ and $W^{pool}(\beta, \overline{\alpha})$, denote respectively the welfare levels associated with a separating and a pooling allocation, when the parental leave rule is set at $\bar{\alpha}$ and the weight assigned to type-1 workers is β. Further, let $W^{sep}(\beta)$ and $W^{pool}(\beta)$ denote the welfare-maximizing allocations under the separating and the pooling regimes, respectively, when the weight assigned to type-1 workers is β. By virtue of our characterization of the welfare-maximizing allocations under the two regimes, for $\beta < \gamma^1$, the optimum for the separating regime is given by a corner solution ($\overline{\alpha} = \alpha^{2*}$) whereas the optimum for the pooling regime is given by an interior solution in which the optimal duration of parental leave satisfies $\overline{\alpha} > \alpha^{2*}$. Thus, it follows that

$$W^{pool}(\beta) > W^{pool}(\beta, \alpha^{2*}) = W^{sep}(\beta, \alpha^{2*}) = W^{sep}(\beta).$$

Moreover, for $\beta = \gamma^1$,

$$W^{pool}(\beta) = W^{pool}(\beta, \alpha^{2*}) = W^{sep}(\beta, \alpha^{2*}) = W^{sep}(\beta).$$

This completes the proof of part (ii).

Part (i) We turn next to prove part (i) by considering the case where $\gamma^1 < \beta \leq 1$. Let $J(\beta) \equiv W^{sep}(\beta) - W^{pool}(\beta)$. Notice that as shown above $J(\gamma^1) = 0$. To complete the proof of part (i) it suffices to show that $J'(\beta) > 0$ for $\beta > \gamma^1$. Using our previous notation, employing the envelope condition and following some algebraic manipulations, one obtains:

$$J'(\beta) = [\hat{U}^1(\overline{\alpha}^{sep}) - \hat{U}^2(\overline{\alpha}^{sep})] - [U^1(\overline{\alpha}^{pool}) - U^2(\overline{\alpha}^{pool})]$$
$$= (\pi^2 - \pi^1)[v(\overline{\alpha}^{pool}) - v(\overline{\alpha}^{sep}),$$

where $\overline{\alpha}^{sep}$ and $\overline{\alpha}^{pool}$ denote the optimal duration of parental leave under the separating and the pooling regimes, respectively. As $\pi^2 > \pi^1$, to complete the proof of part (i) it suffices to show that $v(\overline{\alpha}^{pool}) > v(\overline{\alpha}^{sep})$. By virtue of the strict concavity of v it therefore suffices show that $v'(\overline{\alpha}^{pool}) < v'(\overline{\alpha}^{sep})$. To see this, we employ the first order conditions for the welfare optimum under the two regimes to obtain:

$$v'(\overline{\alpha}^{pool}) = \frac{(\gamma^1\pi^1 + \gamma^2\pi^2)}{[\beta\pi^1 + (1-\beta)\pi^2]} \quad \text{and} \quad v'(\overline{\alpha}^{sep}) = \frac{\gamma^1\pi^1}{\beta\pi^1 + (\gamma^1 - \beta)\pi^2}.$$

We thus need to show that:

$$\frac{\gamma^1\pi^1}{\beta\pi^1 + (\gamma^1 - \beta)\pi^2} > \frac{(\gamma^1\pi^1 + \gamma^2\pi^2)}{[\beta\pi^1 + (1-\beta)\pi^2]}.$$

Re-arranging the left-hand side of the above inequality yields:

$$\frac{(\gamma^1\pi^1 + \gamma^2\pi^2) - \gamma^2\pi^2}{[\beta\pi^1 + (1-\beta)\pi^2] - \gamma^2\pi^2} > \frac{(\gamma^1\pi^1 + \gamma^2\pi^2)}{[\beta\pi^1 + (1-\beta)\pi^2]},$$

which holds if-and-only-if:

$$(\gamma^1\pi^1 + \gamma^2\pi^2) > [\beta\pi^1 + (1-\beta)\pi^2].$$

The latter inequality follows as $\pi^2 > \pi^1$ and $\beta > \gamma^1$. This completes the proof of part (i).

Part (iii) Part (iii) follows immediately, by noticing that the optimum is given by an interior solution in both ranges, characterized in parts (i) and (ii) and recalling that within the ranges of the interior solution the optimal duration under both the separating and the pooling regimes is decreasing with respect to β. This completes the proof.