The Welfare-Enhancing Role of Parental Leave Mandates*

Spencer Bastani† Tomer Blumkin‡ Luca Micheletto§

May 15, 2018

Abstract

A major factor that contributes to persistent gender variation in labor market outcomes is women's traditional role in the household. Child-related absences from work imply that women accumulate less job experience, are more prone to career discontinuities and, hence, suffer a motherhood penalty. We highlight how the gender-driven career/family segmentation of the labor market may create a normative justification for parental leave rules as a means to enhance efficiency in the labor market and alleviate the gender wage gap.

Keywords: parental leave mandates, efficiency, gender wage gap, redistribution

JEL classification: D82, H21, J31, J83

*A previous version of this paper circulated under the title, 'Anti-discrimination legislation and the efficiency-enhancing role of mandatory parental leave'. We are grateful to Dan Anderberg, Catherine Cuff, Thomas Giebe, 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, Linnaeus University, University of Siegen, the CESifo Employment and Social Protection Conference, the International Institute of Public Finance Annual Meeting in Lake Tahoe, the Association for Public Economic Theory Annual Meeting in Paris, the Annual Meeting of the Israeli Economics Association in Tel Aviv, Bocconi University, and the Research Institute for Industrial Economics (IFN) for helpful comments on an earlier draft of the paper. Financial Support from Riksbankens Jubileumfond and the Jan Wallander and Tom Hedelius Foundation is gratefully acknowledged.

†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 voluminous body of evidence documenting gender variation in labor market outcomes, including differences in employment rates, working hours, earnings and job composition (in terms of sector, occupation type and scope). A recent survey by Olivetti and Petrongolo (2016) reviews the existing literature and points out that despite a post-war convergence process, reflecting a host of supply side factors, including, medical advances (availability of birth control), human capital investment (access to higher education) and family-friendly policies (provision of affordable child care services and generous parental leave arrangements), substantial gender differences in pay and employment levels still remain.

A major factor that contributes to persistent gender gaps in labor market performance is parenthood. Women, who traditionally take the lion's 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.[1] Indeed, there is now a large and growing empirical literature documenting the wage penalty associated with motherhood. For instance, for women in the US, the average wage penalty associated with an additional child is around 5%, and persists even when workplace factors and education are controlled for (Waldfogel 1997, Budig and England 2001).

Goldin (2014) argues that workplace flexibility is a key factor in explaining gender wage differences. She discusses multiple dimensions of job flexibility, including the number of hours at work, the precise (particular) times worked, and the predictability and ability to set the work schedule. She further argues that such flexibility is costly for the firm, alluding to mechanisms such as the limited ability of workers with discontinuous work schedules to interact with co-workers and clients, which may hamper the transmission of vital job-related information. This implies that flexible jobs pay less, and workers are faced with a fundamental trade-off between flexibility and compensation. For instance, high skill mothers may compromise by selecting into part-time, low-level, flexible jobs, rather than pursuing a professional challenging career, as documented by Blau and Kahn (2013).

The degree of workplace flexibility granted to workers reflects institutional arrangements in the labor market and prevailing norms but is, to a large extent, shaped by government policy. A notable example is parental leave rules. The latter, taking a broad perspective, 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 additional aspects of workplace

[1] See Bertrand et al. (2010) who focus on workers in the corporate and financial sector. Bertrand et al. 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. Further evidence on the important effects of child-related absences on labor market outcomes is presented by Angelov et al. (2016).

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.

There are large differences across countries in terms of the generosity of parental leave, such as the duration of leave, the level of benefits, job protection features and eligibility (for a comprehensive recent survey see Rossin-Slater 2017). The United States is a country with one of the least generous systems where the flexibility of labor contracts with respect to child-related absences is largely a decision made by employers. The federal Family and Medical Leave Act, which ensures that parents can leave their jobs for 12 weeks and then come back, does not apply to small firms with less than 50 employees. 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.[2]

In this paper we explore a novel normative justification for parental leave rules in the presence of a fundamental gender-driven career/family (compensation/flexibility) conflict faced by workers in the labor market. We consider a benchmark framework where firms are unable to offer distinct contracts to workers differing in their career/family orientation due to asymmetric information, or, by virtue of anti-discrimination legislation that prevents them from doing so. We show that in such a setting, a distortion arises taking the form of an underprovision of workplace flexibility. We then demonstrate that the government can make use of parental leave mandates, as a means to regulate the extent of workplace flexibility, thereby mitigating the distortion and also promoting redistributive goals via reducing the extent of gender pay gaps.

The paper is organized as follows. In section 2 we discuss the related literature, in section 3 we outline our model and present the efficient symmetric information laissez-faire allocation. We then present our benchmark setting with asymmetric information and demonstrate the distortion associated with an underprovision of workplace flexibility. In section 4 we show how introducing parental leave mandates can possibly mitigate this distortion, resulting in a Pareto-improvement. In section 5 we examine the socially desirable parental leave policy and discuss the implications for gender equality. Section 6 presents some extensions of our analysis and, finally, section 7 offers concluding remarks.

2 Related literature

Since the seminal contribution of Rothschild and Stiglitz (1976), minimum coverages have been identified as means to achieve efficiency in insurance contexts where the focus is on asymmetric

[2]A country with one of the world's 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 government subsidized parental leave. Unclaimed days can be saved, and used for parental leave spells up until the child is 8 years old. This is supplemented by generous sick-leave arrangements allowing parents to take up to 120 days off work per year for each sick child under the age of 12. In addition, parents in Sweden have the right to work 75% out of the normal (full-time) weekly working hours until the child is 8 years old.

information regarding heterogeneity in risk types. A general discussion of the social desirability of mandates is provided by Summers (1989) who emphasize the role played by mandates in correcting for externalities, and the connection between mandates and public good provision. In the context of annuities, Eckstein et al. (1985) show that a minimum mandated coverage and the ability to buy insurance beyond the minimum can yield Pareto improvements. In the context of health insurance, mandates have been analyzed by Neudeck and Podczeck (1996), Encinosa (2001), Finkelstein (2004) and McFadden et al. (2015), using a variety of equilibrium concepts. Mandates have also been discussed in the sick-leave literature. For example, Pichler and Ziebarth (2017) discuss how mandated paid sick leave can be desirable on efficiency grounds in a setting where firms have imperfect information about the contagiousness of sick workers present in the workplace.[3]

Our paper differs from the above literature in several ways. We study an averse selection problem deriving from differences in preferences (career-family orientation) which in conjunction with anti-discrimination legislation induce firms to behave *as if* they operate under asymmetric information. We thus take insights from the insurance literature and apply them in a different context where we study the role of parental leave mandates. We shed light on positive aspects (such as motherhood penalties and gender wage gaps arising in our benchmark equilibrium), the efficiency aspect of parental leave as a means to inject missing flexibility in the labor market, as well as normative aspects of an optimal parental leave system that mitigates gender wage gaps. We consider novel aspects such as the combination of mandates and taxation and the possibility to attain a pooling equilibrium through an appropriately chosen parental leave mandate, which has implications for gender equality in the labor market.

By providing a normative justification for parental leave mandates, our paper also relates to a small and recent theoretical literature on the effects of this particular form of government intervention. This includes Barigozzi et al. (2017) and Del Rey et al. (2017). The former contribution emphasizes the interaction between parental leave policy, externalities generated through endogenous social norms concerning child care activities, and career choices; the latter investigates the effects of parental leave on unemployment and wages in a search and matching model of the labor market. More broadly, our paper also relates to a recent strand in the literature that examines the optimal design of public policy in environments with two layers of asymmetric information, between the government and the agents and between private actors in the market (see Stantcheva 2014, Bastani et al. 2015, and Cremer and Roeder 2017). Moreover, our paper relates to Thomas (2018) who analyze the welfare effects of government-mandated maternity leave policies from the perspective that they can change employer's expectations of women's future labor supply, and therefore, the incentives for employers to invest in their workers.[4] Another related paper is Bronson (2015) who constructs and estimates a dynamic structural model

[3]Although they do not model the behavior of firms in their theoretical analysis.

[4]Using data from the US, she shows that such policies increases female employment but decreases the likelihood of women getting promoted.

of marriage, education choices, and life-time labor supply where the "work-family"-flexibility of jobs and educational tracks plays an important role.[5]

3 Model

We consider a simple labor market with an identical number of equally skilled female- and male-workers. Each worker is endowed with a fixed amount of time (normalized to unity) that is allocated between work time and time spent with his/her children (parental leave). Workers differ in their career/family orientation, which is reflected in the probability to take parental leave. We simplify by assuming that workers can be either career-oriented (with a low-probability to take a leave) or family-oriented (with a high-probability to take a leave). We refer to career-oriented workers as type 1 and to family-oriented workers as type 2, and denote their respective fractions in the population by $\gamma^i, i = 1, 2$ where the total population size is normalized to unity, without loss of generality. We let θ^j, $j = m, f$, denote the fraction of family oriented workers among men and women, respectively. We assume realistically that $\theta^f > \theta^m$, reflecting that women exhibit, on average, a stronger family-orientation. This implies that $\gamma^2 = \frac{\theta^f + \theta^m}{2}$ and $\gamma^1 = 1 - \gamma^2$.

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 total consumption (over the unit endowment of time), α denotes the duration of parental leave and π^i denotes the probability that a type i worker will take parental leave (where $\pi^2 > \pi^1$). The function v is assumed to be strictly increasing and strictly concave.

For tractability, we simplify by invoking a quasi-linear specification and by assuming that the probability to take parental leave, π^i, is exogenous. We relax both these assumptions in section 6.3, where we show that these assumptions do not change the qualitative nature of our results.

We assume that the labor market is perfectly competitive. Firms do not observe the family orientation of workers or, alternatively, firms are prevented by some form of anti-discrimination legislation, from tagging workers based on observable characteristics correlated with family orientation, such as gender, age, the number and age of children, or marital status.[6] Thus, firms instead rely on screening employment contracts. A typical labor contract offers the worker a given amount of (monetary) compensation y (which is also equal to the workers' consumption c in the absences of taxes and transfers), and a given duration of parental leave, α. The latter

[5] Using her calibrated model, she performs simulations of different policies, including family-friendly policies such as paid parental leave, part-time work entitlements and subsidized child care, noticing the sometimes ambiguous effects these policies can have on gender equality in the labor market.

[6] As we have decided to focus on gender-related issues, we have refrained from explicitly modeling these other dimensions of heterogeneity which are also likely to be correlated with career/family orientation.

captures the extent of workplace flexibility, which essentially is an option to take a pre-specified amount of time off work, should family circumstances demand it.

Free entry implies that firms break even in expectation. Thus, a firm offering a contract (y^i, α^i) to a type-i worker must satisfy

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

where $\pi^i \alpha^i$ is the total expected time worker i will be away from work and where we have normalized the productivity of the worker per unit of time to 1. The greater the extent of workplace flexibility offered to a worker, the lower is the compensation that he/she receives. Our setup thus captures in a simplistic form a fundamental tension between compensation and flexibility.

The variation in career/family orientation across workers affects both the demand and supply of workplace flexibility. From the workers' perspective, a stronger family orientation is reflected in a higher willingness to pay for additional flexibility (extended parental leave). From the firms' perspective, a family-oriented worker, being more likely to be absent from work, is (in expected terms) less productive than an equally skilled career-oriented counterpart.

3.1 Symmetric information equilibrium

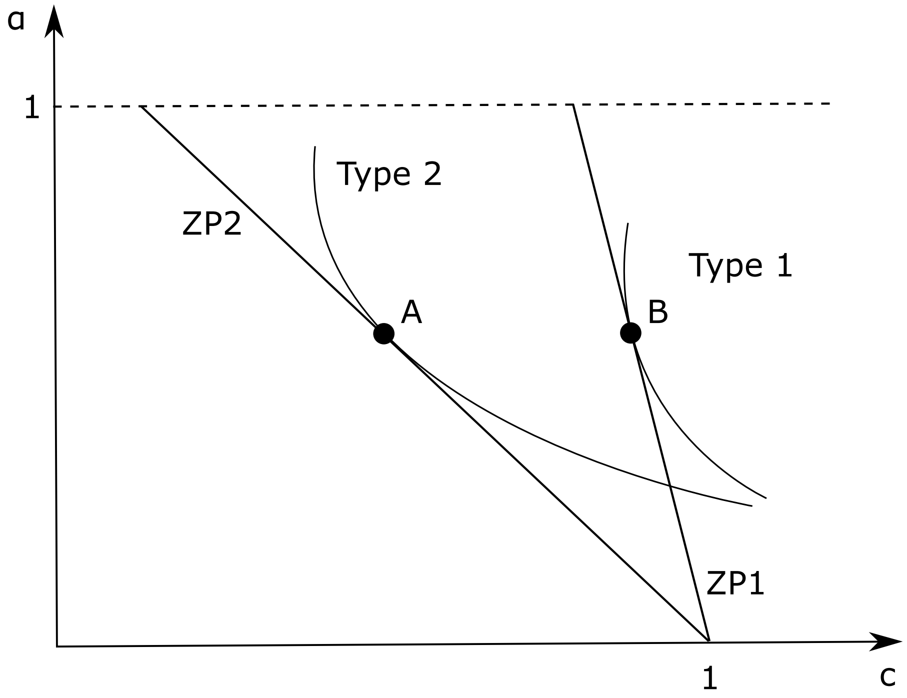
The symmetric information equilibrium reflects a situation where firms are able to observe or infer the family orientation of workers, and, at the same time, are not prevented by law from discriminating between workers. Thus, in the symmetric information equilibrium, the firm can offer distinct contracts to workers of type 1 and type 2. The symmetric information contracts maximize the utility in (1) subject to the budget constraint (2) resulting in an efficient labor market equilibrium satisfying the familiar tangency condition:

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

The optimal contract, for each type of worker ($i = 1, 2$), is given by the solution to a system of two equations: the zero profit condition (budget constraint) in (2) and the tangency 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.

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 $\pi^2\alpha^2 > \pi^1\alpha^1$. The two contracts offer the same degree of flexibility $(\alpha^1 = \alpha^2)$ but family-oriented workers suffer a wage penalty as they have a higher expected workplace absence ($\pi^2\alpha^2 > \pi^1\alpha^1$).

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.



The fact that in the efficient symmetric information allocation, both types of workers are offered the same parental leave is driven by our simplifying assumptions which are made for tractability.

First, we have assumed that both types of workers derive the same utility from child-related absences from work. It is, however, plausible that family-oriented workers would assign a higher value to time spent with their children. Second, we have assumed that the only difference between the two career paths is in the total expected time that workers spend on the job, which is reflected in their output and remuneration. In reality, not only total hours, but also the particular hours matter (i.e. being available for clients and peers at the workplace), suggesting that workers holding family friendly (flexible) jobs are less productive *per unit of time* relative to equally skilled workers who hold career-oriented jobs (see Goldin 2014 and our discussion in the introduction).[7] These assumptions are relaxed in two extensions of our model in section 6 where family-oriented workers, in the efficient symmetric information equilibrium, are offered a longer parental leave than their equally skilled career-oriented counterparts (and, accordingly, suffer an even larger wage penalty).

[7]Notice that becoming less productive per unit of time is likely to be largely determined by having a low degree of workplace attendance in earlier time periods. This is the mechanism of human capital accumulation emphasized in Bronson (2015) and Blundell et al. (2016). Thus, one can interpret (2) as not only reflecting the instantaneous loss in output due to workplace absence, but also the expected productivity losses tomorrow due to a higher workplace absence today.

3.2 Benchmark equilibrium with asymmetric information

We now proceed to consider the realistic case where firms either can not observe or infer the family-orientation of workers, or, are prevented from offering distinct contracts to type 1 and type 2 workers, due to some form of anti-discrimination legislation. This will serve as our benchmark equilibrium.

When firms are not allowed to discriminate directly, they do it indirectly, by offering workers the choice to self-select into two career paths: 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.

The labor market equilbirum 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).[8] Figure 2 illustrates the separating equilibrium, along with the symmetric information equilibrium. Notice that a pooling equilibrium in which both types of workers are offered an identical bundle cannot exist due to 'cream-skimming'. Firms can always break a pooling contract, and derive positive profits, by offering a bundle (with lower flexibility and higher compensation), which would attract career-oriented workers only.[9]

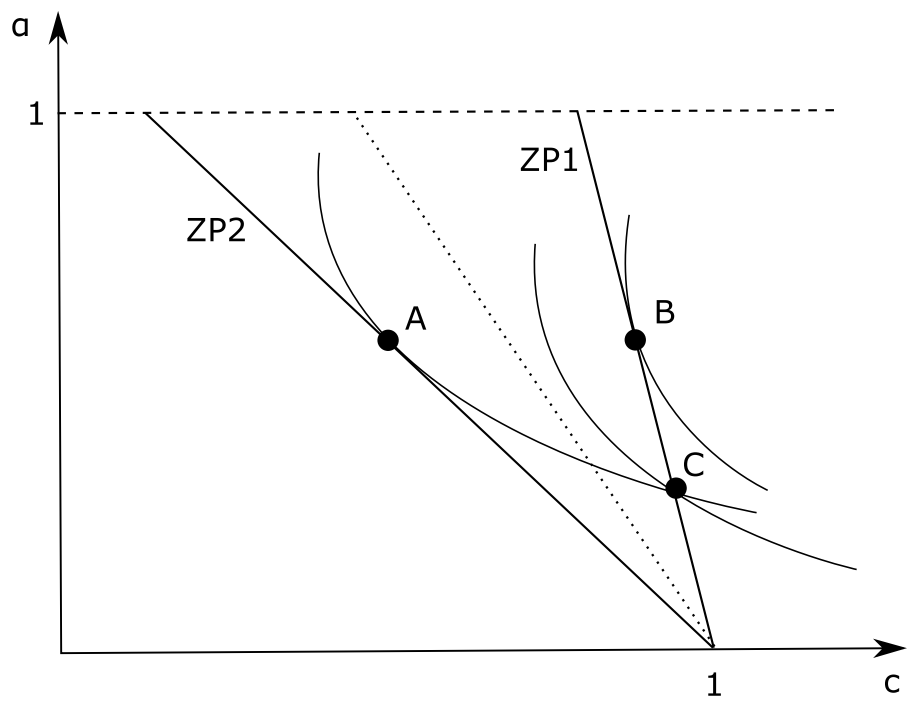
Notice that when the efficient contracts from section 3.1 (points A and B in the figure) are available to both types of workers, each of them will prefer the contract intended for type 1 workers (point B in the figure). Hence, the separating allocation associated with the symmetric information case cannot form an equilibrium, since it is not incentive compatible.

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 asymmetric information. However, type 1 workers must be offered the contract depicted by point 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 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 obtain less parental leave, and correspondingly obtain a higher compensation, than under the symmetric information equilibrium, yielding them a lower level

[8] We follow the notion of equilbrium suggested by the seminal paper by Rothschild and Stiglitz (1976).

[9] 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 dotted 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 6.

Figure 2: Benchmark equilibrium. 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.



of utility.[10]

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

Definition 1 (Benchmark equilibrium)**.** *The benchmark labor market equilibrium 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).*

In order for the separating equilibrium to exist, we need 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. To ensure existence of a separating equilibrium, 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. In figure 2, this assumption is satisfied.

3.3 The gender wage gap in the benchmark equilibrium

While the key focus of this paper is normative issues, there is an important positive aspect of our benchmark equilibrium that we would like to point out, and which relates to the recent debate about the persistent gender wage gaps mentioned in the introduction.

In the benchmark equilibrium, family-oriented workers earn less than career-oriented workers of the same skill level. As we have assumed, realistically, that there is a higher fraction of family-oriented workers among women (i.e. $\theta^f > \theta^m$), it follows that in the benchmark equilibrium, on average, female workers suffer a wage penalty relative to their equally skilled male counterparts. Notice that this is the case even in the symmetric information efficient equilibrium, but the inability of firms to perfectly tag workers further exacerbates the gender pay gap relative to the symmetric information equilibrium. Ironically, to the extent that the inability

[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.

of implementing the symmetric information equilibrium is related to anti-discrimination legislation preventing firms from engaging in tagging, one potential implication would be that anti-discrimination legislation, in and of itself, is counter-productive in promoting equal pay for men and women in the labor market.

Our model is consistent with empirical evidence alluding to the importance of sorting across career tracks as an explanation for gender differences in labor market outcomes, where female workers to a larger extent sort into family-friendly tracks.

4 The efficiency enhancing role of parental leave

We now focus on the consequences of setting a lower bound on the duration of parental leave at a level that is slightly above the amount prescribed, at the benchmark equilibrium, by the contract intended for career-oriented workers. Formally, a binding mandatory parental leave rule, denoted by $\bar{\alpha}$, implies that in equilibrium the following condition has to hold: $\alpha^i \geq \bar{\alpha}; i = 1, 2$, where $\bar{\alpha} > \alpha^{1*}$. As it turns out, the government can use a parental leave mandate to inject the 'missing' flexibility into the labor market, thereby correcting the market failure present in the benchmark equilibrium.[11] In section 4.1 we characterize the labor market equilibrium in the presence of a parental leave mandate. We start by an informal (graphical) description and then provide a formal definition. In section 4.2, we turn to address the normative question regarding the desirability of parental leave mandates on efficiency grounds.

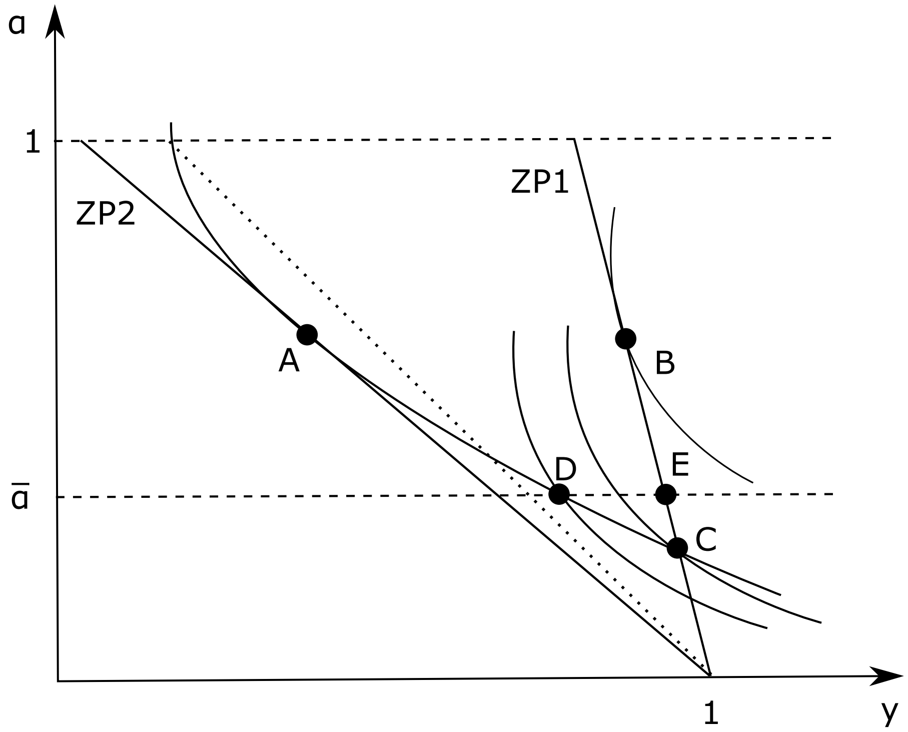
4.1 Equilibrium in the presence of a parental leave mandate

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 properties of the benchmark equilibrium carry over to the separating equilibrium with parental leave.

In figure 3 we present the benchmark equilibrium, illustrated as points A and C in the figure, along with a binding parental leave rule $\alpha = \overline{\alpha}$. The parental leave rule is chosen to be binding, so that it renders the point C infeasible but does not constrain the efficient contract offered to type 2 (point A). The fundamental difference between the benchmark allocation and the one arising in the presence of a parental leave rule is that in the former, the allocation of a type 1 worker is given by the intersection of the indifference curve of type 2 worker (going through his/her equilibrium allocation) and the zero profit line associated with firms hiring type 1 workers (point C in figure 3), whereas in the latter, it is given by the intersection of

[11] In the 'first best' sense, as demonstrated above, the benchmark allocation is inefficient. Our purpose, however, is to examine whether the benchmark equilibrium is second-best inefficient given the policy tools available to the government.

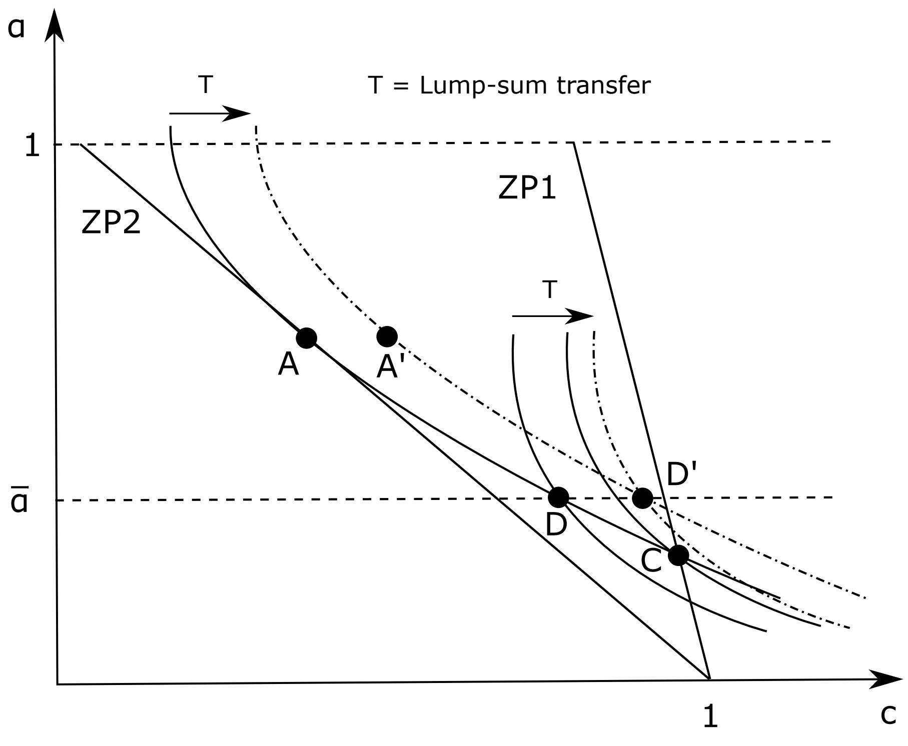
Figure 3: Equilibrium with parental leave. The contract depicted by point C in the figure is no longer feasible due to the presence of the parental leave rule.



the indifference curve of type 2 (going through his/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 them. This is illustrated in figure 3 by virtue of the fact that point D lies below the zero profit line ZP1. The reason firms hiring type 1 workers can derive positive profits in equilibrium is the binding parental leave. The latter prevents a new firm from entering the market and offering a contract with a slightly lower value of α in exchange for a slightly higher compensation, which would attract type 1 workers only, and still maintain non-negative profits. We assume that the government taxes these profits and rebate the tax revenues back to agents in a lump-sum manner. An illustration of an equilibrium with both the parental leave rule and the lump-sum transfer in place can be found in figure 4. Due to quasi-linearity, the lump-sum transfer is reflected in an equal shift of points A and D to the right. The figure illustrates a Pareto-improvement relative to the benchmark allocation, but this does not necessarily need to be the case. We explore this in the next subsection.

Figure 4: Illustration of parental leave + lump-sum transfer. The lump-sum transfer implies (due to quasi-linear utility) a shift to the right of points A and D to, for example, points A' and D', where the distances AA' and DD' are equal. In the example, D' lies to the right of the indifference curve associated with type 1 in the benchmark equilibrium, hence a strict Pareto improvement is achieved.



To formally define the separating equilibrium associated with a parental leave rule, supplemented by pure profits taxation and a (universal) lump-sum transfer, let the profits associated with the contract offered to type 1 workers be denoted by $\sigma > 0$. Total tax revenues associated with this tax on (pure) profits are hence given by $\gamma^1\sigma > 0$. These tax revenues are rebated back to agents in a lump-sum manner. As the population size is normalized to unity, this (universal) lump-sum transfer is also equal to $\gamma^1\sigma$.

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 allocation* $(c^1(\sigma^*), \overline{\alpha})$ *and* $(c^2(\sigma^*), \alpha^2(\sigma^*))$ *where* σ^* *is the solution to:*

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

and

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

$$c^1(\sigma) = 1 - \pi^1\overline{\alpha} - \sigma + \gamma^1\sigma. \tag{6}$$

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 $\gamma^1\sigma$. The consumption of type 1 agents, given by condition (6), 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, $\alpha^2(\sigma) = \alpha^{2*}$ (where α^{2*} is the duration of parental leave for a type 2 agent in the benchmark allocation; see definition 1), hence condition (4) simplifies to

$$1 - \pi^2\alpha^{2*} + \gamma^1\sigma + \pi^2 v(\alpha^{2*}) = 1 - \pi^1\overline{\alpha} - \gamma^2\sigma + \pi^2 v(\overline{\alpha}). \tag{7}$$

In addition to the simplified condition given in (7), 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, given the tax system in place. Formally, the following condition has to hold:

$$\max_{\alpha \geqslant \overline{\alpha}} 1 - \alpha \sum \gamma^i\pi^i + \gamma^1\sigma + \pi^1 v(\alpha) \leq 1 - \pi^1\overline{\alpha} - \gamma^2\sigma + \pi^1 v(\overline{\alpha}). \tag{8}$$

The LHS of the inequality in (8) describes the utility associated with the pooling contract, along the zero-profit line, given the lump-sum transfer $\gamma^1\sigma$. The RHS is the separating allocation associated with type 1, as characterized above.[12]

[12]Notice that condition (8) is not equivalent to assumption 1, invoked to imply that a separating equilibrium

One can show that the equilibrium characterized by the equation in definition 2 is well-defined, namely, 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 (7). 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 $\frac{\partial}{\partial \overline{\alpha}}[1 - \pi^1 \overline{\alpha} + \pi^2 v(\overline{\alpha})] > 0$, 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 (7) 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 (7) 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 (7). As the RHS is strictly decreasing in σ and the LHS is strictly increasing in σ, the solution is unique.

4.2 When is parental leave efficiency-enhancing?

We now proceed to discuss when the composite reform described in the previous section is Pareto-improving. One may first notice that type 2 workers are unambiguously made better off as they obtain the same labor contract as in the benchmark equilibrium, but in addition receive a lump-sum transfer from the government. Turning next to type 1 workers, there are two conflicting forces at play that determine whether type 1 workers become better off from the reform and, ultimately, whether a Pareto-improvement is attainable. First, introducing a parental leave mandate induces firms hiring type 1 workers to offer a higher level of α. This shifts the type 1 contract in the direction of the first best contract, mitigates the distortion associated with the benchmark allocation and makes type 1 workers better off. Second, the combination of the confiscatory tax levied on the profits of firms hiring type 1 workers, and the universal lump sum transfer, implies that type 1 workers are effectively paying a tax. This implies that their consumption level is shifted to the left of the zero profit condition (ZP1), making them worse off. Without further restrictions, we only know that the allocation of type 1 workers, after the composite reform, lies along the line $\alpha = \overline{\alpha}$ in figure 3, between points D and E, which may or may not entail that type 1 workers become better off.

The following proposition provides a necessary and sufficient condition for a Pareto-improvement relative to the benchmark equilibrium.

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

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

where α^{1*} is associated with the separating benchmark equilibrium.

exists, due to the tax system in place. Nonetheless, assumption 1 implies that (8) is satisfied, by continuity, provided that the degree of cross-subsidization induced by imposing the binding parental leave rule is sufficiently small.

Proof See appendix A. □

The condition in Proposition 1 is expressed in terms of the features 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 an improvement to be feasible. When the extent of induced cross-subsidization is small (γ^2 is small) and/or the distortion is large (α^{1*} is small) the case for parental leave becomes stronger. 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.[13] The effect of differences in π on the likelihood to obtain a Pareto-improvement is instead generally ambiguous.[14]

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. 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 slightly higher compensation, thereby attracting type 1 workers (who are in expectation more productive) 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 necessary and sufficient to achieve a Pareto improvement.[15] In a numerical example in appendix C we demonstrate that it is possible to simultaneously satisfy the existence condition on page 10 and the condition for Pareto improvement (9) for a wide range of parameter values.

Notice that the cross-subsidization from career- towards family-oriented workers associated with our composite parental leave reform, serves to reduce the gender wage gap, thereby promoting redistribute goals. We explore this in the next section. However, before moving on to this topic, we briefly discuss the role of subsidized parental leave.

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

[14]This is shown in appendix B where we also resolve this ambiguity in a numerical example given certain parametric assumptions.

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

4.3 Subsidized parental leave

As mentioned in the introduction, in most OECD countries (the US being the exception) the government is subsidizing the child-related absences from work that are mandated by law.[16]

Suppose that in contrast to section 4.1, where we assumed that revenues from profits taxation are rebated in a lump-sum fashion across the board, benefits are paid out as a function of the time spent on leave (as is the case in many countries). We refer to this as a "subsidized parental leave system". In appendix F we show that there is no normative justification, at least on efficiency grounds, for the commonly observed pattern of subsidized leave. The reason for this derives from the fact that in the benchmark equilibrium the incentive constraint associated with family-oriented workers is binding. In order to expand the set of parameters for which a Pareto improvement can be obtained, one has to use policy tools that mitigate this incentive compatibility constraint by rendering it less attractive for family-oriented workers to mimic their career-oriented counterparts. In appendix F, we show that a subsidized parental leave system can obtain a Pareto-improvement for a smaller set of parameters relative to a parental leave system with uniform benefits because subsidized PL is more attractive to family oriented workers and hence renders mimicking more attractive.[17]

5 The socially desirable duration of parental leave

In section 4 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?

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 $\alpha^{1*} < \alpha^{2*}$.

To analyze the optimal duration of parental leave one must acknowledge that, depending on the value of $\overline{\alpha}$, the government might be implementing either a separating or pooling labor market allocation. 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. The separating equilibrium in the presence of our composite parental leave policy was formally described in definition 2 on p. 14. For completeness, we provide below a formal definition of the pooling equilibrium in the presence of parental leave. Notice that, as there are no expected profits, there

[16] According to an analysis by the International Labor Organization of the United Nations, 74 out of 167 countries with available data provide maternity benefits that amount to at least two-thirds of a woman's previous earnings for at least 14 weeks. Out of them, 61 countries provide 100 percent of prior earnings for 14 weeks (see Rossin-Slater 2017).

[17] In a working paper version of this paper (Bastani et al. 2016, appendix C) we show that allowing to tax children (rather than providing benefits) can expand the set of parameters for which a Pareto-improvement can be obtained. The potentially welfare enhancing role of taxing children has previously been emphasized by Balestrino et al. (2002) and Cigno and Pettini (2002).

are no taxes and transfers associated with the pooling regime.

Definition 3. *The pooling equilibrium associated with a parental leave rule is given by the allocation* $(\hat{c}, \bar{\alpha})$ *where*

$$\hat{c} = 1 - \overline{\alpha}(\gamma^1\pi^1 + \gamma^2\pi^2).$$

We begin by characterizing the second best Pareto frontier, letting the policymaker maximize a weighted average of the utilities of both types of agents, and then we proceed to discuss the implications for gender equality.

5.1 Characterization of the second best Pareto frontier and implications for gender equality

The social maximization problem is defined as follows:

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

where $U^i(\overline{\alpha}, j)$ 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, formally described in definitions 2 and 3) 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.

The following proposition characterizes the second best Pareto frontier.

Proposition 2 (Characterization of the Social Optimum)**.**

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

(ii) 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 D. □

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 by a pooling equilibrium, when the weight attached to career-oriented workers is relatively low. Furthermore, the optimal duration of parental leave is increasing with respect to the weight assigned to family oriented workers (type 2).[18]

[18] 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.

There are two considerations that the government takes into account in the welfare maximization. The first is the efficiency-enhancing role of the parental leave mandate to correct the distortion associated with the benchmark equilibrium. The second concerns the redistributive role played by the parental leave policy due to the induced cross-subsidization from type 1 to type 2 workers. How the government values this redistribution is captured by β, which is the weight attached to type 1 (career-oriented) workers. A higher β is thus reflecting a stronger bias of the government in favor of type 1 workers, and vice versa.

Consider first the case when $\beta = \gamma^1$, which implies that the weight attached to each type of worker is exactly equal to its share in the population, namely, there is no bias in favor of either worker. In such a case, by virtue of quasi-linearity, the only consideration at play is the efficiency-enhancing role of parental leave, and the resulting social optimum implements the efficient allocation of parental leave associated with the symmetric information equilibrium.[19]

Consider next the case when the government is biased in favor of career-oriented workers ($\beta > \gamma^1$). In this case, the social optimum is characterized by a separating equilibrium where a binding parental leave rule is implemented, but the duration is shorter than the duration associated with the efficient symmetric information equilibrium. The simple reason for this is the cross subsidization from career- to family- oriented workers, induced by extending the duration of parental leave. Thus, some efficiency is sacrificed in order to promote redistributive goals, which in this case suggests redistributing from type 2 to type 1 workers, via reducing the extent of cross-subsidization (discussed above).

Finally, and perhaps most realistically, in the case where the government places a higher weight on family-oriented workers ($\beta < \gamma^1$) the social optimum is given by a pooling allocation where the duration of parental leave strictly exceeds the (common) duration of parental leave associated with the efficient symmetric information equilibrium. In this case, the government extends the duration of the parental leave mandate beyond the point of efficiency, as by doing so the government enhances the degree of cross-subsidization and thereby promotes redistribution in favor of family oriented workers. As in the previous case, albeit in an opposite direction, the government sacrifices some efficiency in favor of redistribution. Notice that for any duration of parental leave shorter than the efficient level, in the case where $\beta < \gamma^1$, an extension of parental leave promotes both efficiency and redistribution at the same time. This is the reason why a separating equilibrium is never optimal for this case, and the optimum is given by a pooling equilibrium. In particular, it implies that it is always optimal to fully eliminate the pay differences between the two types of workers whenever $\beta < \gamma^1$.[20]

[19]Notice that relative to the symmetric information equilibrium, there is a difference in the consumption allocation due to the induced cross-subsidization in the parental leave regime. However, due to quasi-linearity, the government is indifferent to how consumption is divided between the two types of workers in this case.

[20]The result that gender wage gaps are fully eliminated, is driven by the fact that in the symmetric information efficient allocation, both types of workers are prescribed the same duration of parental leave. Modifying our framework, by assuming either that family-oriented workers derive a higher utility from being absent from work than their career-oriented counterparts, or, that workers choosing the career-track are more productive than equally skilled workers opting for the family track, will change this result. In particular, it will imply that in a symmetric

The case where the government places a higher weight on family-oriented workers highlights the idea of using family-friendly policies, in particular parental leave mandates, to promote gender equality. Due to the correlation between gender and family-orientation, parental leave becomes an indirect channel through which gender equality can be promoted. In a way, the government is using family-orientation as a tag for gender. This is related to the literature on 'color-blind' affirmative action, which refers to government policies that indirectly target disadvantaged groups in cases where direct targeting would be politically controversial.[21] In this way, the government is sacrificing some efficiency to render policies more socially acceptable. In our setting, the government promotes gender equality via the cross-subsidization induced by the gender-neutral parental leave rule.

6 Extensions

At the end of section 3.1, we briefly discussed the property of the baseline model that in the symmetric information equilibrium, both agents receive the same duration of parental leave α. In this section, we present two simple extensions where family-oriented workers obtain a longer duration of parental leave in the symmetric information equilibrium. In section 6.1, we consider a model where family-oriented workers attach a higher utility to parental leave spells and π is endogenously chosen. In section 6.2 we allow for two different job types, "normal" and "flexible" jobs, where the latter are associated with a wage penalty but provide the benefits of flexibility, valued more highly by family-oriented workers. In both these extensions, when mimicking their career-oriented counterparts, family-oriented workers will receive a shorter duration of parental leave in exchange for a higher compensation. It is hence, in contrast to the baseline model, not obvious that the symmetric information equilibrium is incentive-incompatible. We show, for both extensions, that there is a non-negligible set of parameters for which the incentive constraint is binding and that our main results for the baseline model are is robust to these extensions, relying on a continuity argument. Finally, in section 6.3 we provide a generalization of the model in section 6.1 allowing the utility of consumption to be nonlinear. As the generalized model does not approach our baseline model in the limit, we cannot rely on continuity considerations. Hence, we provide instead for this model a full characterization and generalization of the necessary and sufficient condition for Pareto improvement in Proposition 1.

information efficient allocation, contracts offered to family-oriented workers will prescribe a longer duration of parental leave, relative to the contracts offered to career-oriented workers. In this case, there will be a threshold $\overline{\beta}$, satisfying $0 < \overline{\beta} < \gamma^1$, such that for β exceeding this threshold, the social optimum will be given by a separating allocation, whereas for β below this threshold, a pooling allocation will be socially optimal. When a separating allocation is socially optimal, the parental leave policy would mitigate rather than fully eliminate the gender wage gap.

[21] See Chan and Eyster (2003) and Fryer et al. (2008).

6.1 Endogenous fertility and quasi-linear utility

In this section, we allow the fertility of agents (described by the parameter π) to be endogenously chosen. To facilitate the interpretation we will henceforth refer to π as the (expected) number of children (capturing the non-deterministic feature of fertility).

In the baseline model we assumed that agents differed with respect to some exogenous π. Here, instead, we assume that family-oriented workers attach a higher utility to parental leave spells (described by the parameter k below). This will generally imply that family-oriented workers will take a greater *number of leaves* in expectation (determined by the endogenous π), but also that their parental leave duration will be longer.

In the symmetric information equilibrium, the contract offered to family-oriented workers will feature, relative to career-oriented workers, a longer parental leave duration. This implies that when mimicking their career-oriented counterparts, family-oriented workers will receive a shorter duration of parental leave in exchange for a higher compensation. It is hence not obvious that the incentive constraint is binding, which is necessary for there to be an efficiency-enhancing role for parental leave.

Below, we outline a model along the lines above and show that one can find parameters of the model such that the incentive constraint is binding. It will also become clear that our baseline model is obtained by letting the differences in preferences be sufficiently small, while still large enough to generate a discrete difference in the π of the two agents. In this way, we are providing a microeconomic foundation for the baseline model analyzed in section 3. In section 6.3 below, we further generalize the model, allowing for nonlinear utility of consumption.

We assume individuals solve the following problem:

$$\max_{\pi^i \in [\underline{\pi}, \bar{\pi}]} [y^i - d\pi^i + \pi^i k^i v(\alpha^i)]$$

where $\bar{\pi} > \underline{\pi} > 0$ and d denotes the costs associated with having a child (for example, the amount spent on child care services). The parameter k^i is the value assigned by a type i worker to parental leave $(i = 1, 2)$. We assume that $k^2 > k^1$ and without loss of generality that $k^2 = k$ and $k^1 = 1$.[22] The solution to the above problem is given by:

$$\pi^i(y^i, \alpha^i) = \begin{cases} \underline{\pi}, & \text{if} \quad k^i v(\alpha^i) \le d \\ \bar{\pi}, & \text{if} \quad k^i v(\alpha^i) > d \end{cases}$$

where we additionally impose that $\pi^i(y^i, \alpha^i) = \underline{\pi}$ when the individual is indifferent between $\underline{\pi}$ and $\bar{\pi}$. We focus on labor market equilibria where type 1 and type 2 workers choose different π.

[22] For tractability, we focus on the quasi-linear case. This is without loss of generality for the purpose of showing that the incentive constraint binds. The reason is that with nonlinear utility of consumption, there would be an income effect working in the direction of making the difference in α between the two types of agents less pronounced in the symmetric information equilibrium. Hence, having nonlinear utility of consumption would just make it more likely that the incentive constraint is binding.

This happens when:

$$v(\alpha^1) \le d \quad \text{and} \quad kv(\alpha^2) > d. \tag{10}$$

Moreover, we assume that a type 2 behaving as a mimicker chooses $\bar{\pi}$, which requires that:

$$kv(\alpha^1) > d. \tag{11}$$

Combining (11) with (10), exploiting that $\alpha^2 > \alpha^1$ (which will be verified below), yields

$$d/k < v(\alpha^1) \le d. \tag{12}$$

For simplicity, we will assume that $d = v(\alpha^1)$, which ensures that this condition is always satisfied as $k > 1$.

6.1.1 Incentive Incompatibility of the Symmetric Information Equilibrium

The symmetric information contracts (y^{1*}, α^{1*}) and (y^{2*}, α^{2*}) solve the following maximization problem:

$$\max_{\alpha} \left\{1 - \pi^i \alpha - d\pi^i + \pi^i k^i v(\alpha^i)\right\}$$
$$y^i = 1 - \pi^i \alpha^i$$

for $i = 1, 2$ where $\pi^1 = \underline{\pi}$ and $\pi^2 = \bar{\pi}$, provided condition (12) holds. The FOC is:

$$v'(\alpha^i) = 1/k^i, \quad i = 1, 2. \tag{13}$$

The incentive constraint is binding iff:

$$y^{1*} - d\bar{\pi} + \bar{\pi} k v(\alpha^{1*}) > y^{2*} - d\bar{\pi} + \bar{\pi} k v(\alpha^{2*})$$

or upon re-arrangement:

$$y^{1*} - y^{2*} + \bar{\pi} k [v(\alpha^{1*}) - v(\alpha^{2*})] > 0. \tag{14}$$

Condition (14) is a necessary and sufficient condition for incentive-incompatibility. Notice that

$$y^{1*} = 1 - \underline{\pi} \alpha^{1*}$$
$$y^{2*} = 1 - \bar{\pi} \alpha^{2*}.$$

Hence, since $\alpha^{1*} < \alpha^{2*}$ by virtue of (13) as $k^2 = k > 1 = k^1$ we have that $y^{1*} > y^{2*}$. Thus, the first term of (14) is positive whereas the second term is negative, leaving the overall sign of the LHS of (14) ambiguous. However, for values of k sufficiently close to 1, the difference between α^{1*} and α^{2*} will be small (due to Eq. 13) whereas the term $y^{1*} - y^{2*}$ will be positive and bounded away from zero since $\bar{\pi} > \underline{\pi}$ and given the fact that we have set $d = v(\alpha^{1*})$. Thus, k can always be chosen close enough to 1 such that condition (14) is satisfied implying that the symmetric information equilibrium is incentive-incompatible. Moreover, in this case a Pareto improvement is, by continuity, achievable as long as the condition in Proposition 1 is satisfied.

6.2 A simple model with productivity differences

Suppose π is again exogenous, as in the baseline model, but that contracts now are characterized by an additional explicit dimension of flexibility. We assume firms offer contracts $(y^1_\delta, \alpha^1_\delta, \delta)$ and $(y^2_\delta, \alpha^2_\delta, \delta)$ where $\delta \in \{0, 1\}$ is an indicator for flexibility. We assume family-oriented workers value flexibility, whereas career-oriented workers do not. There are thus three potential contracts in this economy, $(y^1, \alpha^1, 0)$ and $(y^2_0, \alpha^2_0, 0)$, and $(y^2_1, \alpha^2_1, 1)$. Furthermore, family-oriented workers obtain a benefit equal to $\bar{U} > 0$ from the flexible job, but suffer a wage penalty implying that the hourly compensation is reduced from unity to $1 - m$ where $0 < m < 1$. Jobs with $\delta = 1$ are flexible in the sense that they allow for e.g., non-standard working hours, and captures that not being at work when others are, or not being available for clients etc. can have a negative impact on productivity.

Focusing on the relevant case where the binding incentive constraint is that linking type 2 to type 1 workers, the utilities that we will need to evaluate are:

$$U^1(c^1, \alpha^1, 0) = c^1 + \pi^1 v(\alpha^1) \tag{15}$$
$$U^2(c^1, \alpha^1, 0) = c^1 + \pi^2 v(\alpha^1) \tag{16}$$
$$U^2(c^2_0, \alpha^2_0, 0) = c^2_0 + \pi^2 v(\alpha^2_0) \tag{17}$$
$$U^2(c^2_1, \alpha^2_1, 1) = c^2_1 + \pi^2 v(\alpha^2_1) + \bar{U} \tag{18}$$

Free entry implies that firms break even in expectation. Thus, the compensation associated with each of the three jobs must satisfy:

$$y^1 = 1 - \pi^1 \alpha^1 \tag{19}$$
$$y^2_0 = (1 - \pi^2 \alpha^2_0) \tag{20}$$
$$y^2_1 = (1 - m)(1 - \pi^2 \alpha^2_1) \tag{21}$$

In the absence of taxes and transfers, $y^1 = c^1$,$y^2_j = c^2_j$ $j = 0, 1$.

6.2.1 Symmetric information equilibrium

In the symmetric equilibrium, the utility of each type of worker is maximized subject to the relevant zero-profit condition/budget constraint. Insert (20)–(21) into (17)–(18) and define:

$$\alpha_1^{2*} = \underset{\alpha_1^2}{\text{argmax}} \left\{ (1-m)(1-\pi^2\alpha_1^2) + \pi^2 v(\alpha_1^2) + \bar{U} \right\}.$$
$$\alpha_0^{2*} = \underset{\alpha_0^2}{\text{argmax}} \left\{ (1-\pi^2\alpha_0^2) + \pi^2 v(\alpha_0^2) \right\}.$$

The symmetric information equilibrium is given by the two contracts $(c^{1*}, \alpha^{1*}, 0)$ and $(c^{2*}, \alpha^{2*}, \delta^*)$ satisfying:

$$\alpha^{1*} = \underset{\alpha^1}{\text{argmax}} \left\{ 1 - \pi^1\alpha^1 + \pi^1 v(\alpha^1) \right\}$$
$$c^{1*} = 1 - \pi^1\alpha^{1*}$$
$$\delta^* = \underset{\delta}{\text{argmax}} \left\{ (1-\delta m)(1-\pi^2\alpha_\delta^{2*}) + \pi^2 v(\alpha_\delta^{2*}) + \delta\bar{U} \right\}$$
$$\alpha^{2*} = \alpha_{\delta^*}^{2*}$$
$$c^{2*} = (1-\delta^* m)(1-\pi^2\alpha_{\delta^*}^2).$$

Notice that the maximization above implies the following FOC:

$$v'(\alpha^{1*}) = 1 \tag{22}$$
$$v'(\alpha^{2*}) = 1 - \delta^* m. \tag{23}$$

Depending on the parameters m and $\bar{U}$, family oriented workers will either be offered the flexible or non-flexible contract. If the benefit from flexibility $\bar{U}$ is sufficiently large relative to the wage penalty m, family-oriented workers will always choose the flexible contract. However, in this case, it is not obvious that the incentive constraint is binding. We need to show that there is some non-empty set $\mathcal{X}$ such that when $(m, \bar{U}) \in \mathcal{X}$, family-oriented workers choose the flexible contract, and the incentive constraint is binding (namely, type 2 workers strictly prefer the symmetric information equilibrium bundle of type 1 workers over their own bundle).

6.2.2 Incentive Incompatibility of the Symmetric Information Equilibrium

A family-oriented worker is at least as well off obtaining the flexible contract as the non-flexible contract, namely $\delta^* = 1$, if the following condition holds:

$$(1-\pi^2\alpha_0^{2*}) + \pi^2 v(\alpha_0^{2*}) \leq (1-m)(1-\pi^2\alpha_1^{2*}) + \pi^2 v(\alpha_1^{2*}) + \bar{U} \tag{24}$$

The incentive constraint is binding if the following condition holds:

$$(1-\pi^1\alpha^{1*})+\pi^2 v(\alpha^{1*}) > (1-m)(1-\pi^2\alpha_1^{2*})+\pi^2 v(\alpha_1^{2*})+\bar{U}. \tag{25}$$

Notice that from (22) and (23) we have that α^{1*}, α_0^{2*}, and α_1^{2*} satisfy $v'(\alpha^{1*})=1$, $v'(\alpha_0^{2*})=1$, $v'(\alpha_1^{2*})=1-m$. Thus, we have that $\alpha^{1*}=\alpha_0^{2*}<\alpha_1^{2*}=\alpha^{2*}$. Hence, removing unnecessary subscripts, we combine (24) and (25) to obtain the following condition:

$$\begin{aligned}(1-\pi^2\alpha^{1*})+\pi^2 v(\alpha^{1*}) &\le (1-m)(1-\pi^2\alpha^{2*})+\pi^2 v(\alpha^{2*})+\bar{U} \\ &< (1-\pi^1\alpha^{1*})+\pi^2 v(\alpha^{1*})\end{aligned} \tag{26}$$

The left side of this inequality is the utility associated with the non-flexible contract. Picking the non-flexible contract implies a lower duration of parental leave and a higher consumption (due to the fact that no wage penalty is suffered, but also due to the higher workplace presence). The RHS is the utility obtained when mimicking, which is always higher than the LHS because the contract associated with career-oriented workers offers a higher consumption (since $\pi^2>\pi^1$).

Subtracting $(1-\pi^2\alpha^{1*})+\pi^2 v(\alpha^{1*})$ from (26) and re-arranging yields:

$$0 \le (1-m)(1-\pi^2\alpha^{2*})+\pi^2 v(\alpha^{2*})+\bar{U}-\Big((1-\pi^2\alpha^{1*})+\pi^2 v(\alpha^{1*})\Big) < \alpha^{1*}(\pi^2-\pi^1)$$

or after some manipulations:

$$0 \le \pi^2\Big((v(\alpha^{2*})-v(\alpha^{1*}))-(\alpha^{2*}-\alpha^{1*})\Big)-m(1-\pi^2\alpha^{2*})+\bar{U} < \alpha^{1*}(\pi^2-\pi^1). \tag{27}$$

Notice that $-m(1-\pi^2\alpha^{2*})\le 0$ and:

$$\pi^2\Big((v(\alpha^{2*})-v(\alpha^{1*}))-(\alpha^{2*}-\alpha^{1*})\Big)\le 0 \tag{28}$$

since $\alpha^{2*}>\alpha^{1*}$, v is concave and $v'(\alpha^{1*})=1$. Hence, for any equilibrium, one can choose $\bar{U}=-\Big(\pi^2\Big((v(\alpha^{2*})-v(\alpha^{1*}))-(\alpha^{2*}-\alpha^{1*})\Big)-m(1-\pi^2\alpha^{2*})\Big)+\epsilon$, where $0<\epsilon<\alpha^{1*}(\pi^2-\pi^1)$ to ensure that the condition in (27) is satisfied.

The proof above shows that for *any* m, one can always choose $\bar{U}>0$ appropriately so that family-oriented workers prefer the flexible job, and the incentive constraint is binding. Notice that our benchmark model is obtained when $m\to 0$ and $\bar{U}\to 0$, hence, by continuity, the necessary and sufficient condition for Pareto-improvement in Proposition 1 still applies, provided m and $\bar{U}$ are sufficiently close to zero.

6.3 Endogenous fertility and non-linear utility of consumption

In this section we generalize the model in section 6.1 by assuming that the utility from consumption is strictly concave (rather than being linear). To facilitate the interpretation, we again refer to π as the (expected) number of children.

The preference of a typical household is represented by the following utility function:

$$U(c,\alpha,\pi,k) = u(c) + \pi k v(\alpha), \tag{29}$$

where u and v are strictly increasing and strictly concave, c denotes consumption, α denotes the duration of parental leave, and k measures the family orientation of the worker. To be consistent with our baseline setup, we assume that $k^2 > k^1$, reflecting the fact that type 2 workers exhibit a stronger family orientation than their type 1 counterparts. We assume that π is endogenously determined by the worker who is seeking to maximize the utility in (29), given the labor contract offered by the firm, (y,α), and subject to the following budget constraint:

$$y - \pi d = c, \tag{30}$$

where d denotes the costs associated with having a child (for example, the amount spent on child care services).

We start by characterizing the symmetric information equilibrium where firms offer distinct contracts to each type of worker (either by observing the type or by tagging based on observable attributes correlated with the worker's family orientation).

6.3.1 The Symmetric Information Case

In the symmetric information case type-i workers; $i = 1,2$, will be offered the contract (y^i,α^i)which solves the following constrained maximization program:

$$\max_{y^i,\alpha^i,\eta}\left[\max_{\pi^i}[u(y^i - d\pi^i) + \pi^i k^i v(\alpha^i)] + \eta(1 - \pi^i\alpha^i - y^i)\right], \tag{31}$$

where η denotes the Lagrange multiplier associated with the zero profit condition. Notice that the program given in (31) is a nested maximization program in which the contract offered by the firm maximizes the utility of the worker subject to the zero profit condition, taking into account the worker's optimal response (determining the optimal number of children). In what follows we will focus on the first-order conditions, assuming the second-order conditions are satisfied.

Formulating the first order conditions with respect to y and α, employing the worker's envelope condition, yields:

$$u'(c^i) - \eta\left(1 + \alpha^i\frac{\partial\pi^i}{\partial y^i}\right) = 0, \tag{32}$$

$$\pi^i k^i v'(\alpha^i) - \eta\left(\pi^i + \alpha^i\frac{\partial\pi^i}{\partial\alpha^i}\right) = 0. \tag{33}$$

where $c^i \equiv y^i - d\pi^i$ and $\pi^i = \pi^i(y^i,\alpha^i)$ denotes the optimal (expected) number of children chosen by a type-i worker satisfying the first-order condition:

$$du'(c^i) = k^i v(\alpha^i). \tag{34}$$

Substituting for η from (32) into (33) yields upon re-arrangement:

$$\frac{\pi^i k^i v'(\alpha^i)}{u'(c^i)} = \frac{\pi^i + \alpha^i \frac{\partial \pi^i}{\partial \alpha^i}}{1 + \alpha^i \frac{\partial \pi^i}{\partial y^i}}. \tag{35}$$

Substituting for $u'(c^i)$ from (34) into (35) and re-arranging yields:

$$\frac{\pi^i d v'(\alpha^i)}{v(\alpha^i)} = \frac{\pi^i + \alpha^i \frac{\partial \pi^i}{\partial \alpha^i}}{1 + \alpha^i \frac{\partial \pi^i}{\partial y^i}}. \tag{36}$$

Condition (36) is, in similarity to the baseline model (where π is fixed and utility is quasi-linear), a tangency condition between the worker's indifference curve and the zero profit condition of the firm in the (α, y)-plane.

6.3.2 The Asymmetric Information Case

Assume now that the firm cannot distinguish between workers with different career-family orientation and/or are prevented by anti-discrimination law from engaging in tagging. To abbreviate notation, we let $k^1 = 1$ and $k^2 = k > 1$. In the asymmetric information case, the labor contract offered to type 2 (family-oriented) workers will remain as in the symmetric information regime and hence will be given by the solution to the constrained optimization program in (31). Denote by (y^{2*}, α^{2*}) the optimal contract offered to type 2 workers and by π^{2*} the associated optimal (expected) number of children chosen by type 2 workers characterized by the first-order conditions in (32), (33) and (34). Further denote by U^{2*} the utility level associated with the labor contract offered to type 2 workers in equilibrium, formally given by:

$$U^{2*} = u(y^{2*} - d\pi^{2*}) + \pi^{2*} k v(\alpha^{2*}). \tag{37}$$

The labor contract offered to type 1 (career-oriented) workers will be given by the solution to the following constrained optimization program:

$$\max_{y^1, \alpha^1, \lambda, \mu} \Big[\max_{\pi^1} [u(y^1 - d\pi^1) + \pi^1 v(\alpha^1)] + \lambda(1 - \pi^1 \alpha^1 - y^1) + \mu \Big(U^{2*} - \max_{\hat{\pi}^2} [u(y^1 - d\hat{\pi}^2) + \hat{\pi}^2 k v(\alpha^1)] \Big) \Big], \tag{38}$$

where λ and μ denote the Lagrange multipliers associated, correspondingly, with the type 1 zero profit condition and the type 2 incentive compatibility constraint.

Notice the difference between the maximization programs in (31) and (38). In the latter case an additional incentive compatibility constraint is introduced to ensure that type 2 workers will refrain from mimicking their type 1 counterparts. In what follows, we assume that the incentive constraint is binding, namely, the symmetric equilibrium separating allocation is not incentive compatible.

Formulating the first order conditions with respect to y^1 and α^1, employing the worker's envelope conditions, yields:

$$u'(c^1) - \mu u'(\hat{c}^1) - \lambda\left(1 + \alpha^1 \frac{\partial \pi^1}{\partial y^1}\right) = 0, \tag{39}$$

$$\pi^1 v'(\alpha^1) - \mu \hat{\pi}^2 k v'(\alpha^1) - \lambda\left(\pi^1 + \alpha^1 \frac{\partial \pi^1}{\partial \alpha^1}\right) = 0, \tag{40}$$

where $\pi^1 = \pi^1(y^1, \alpha^1)$ denotes the optimal (expected) number of children for type 1 workers, and where $\hat{c}^1 \equiv y^1 - d\hat{\pi}^2$ and $\hat{\pi}^2 = \hat{\pi}^2(y^1, \alpha^1)$ denote, respectively, the consumption level and the optimal (expected) number of children chosen by type 2 workers mimicking their type 1 counterparts, satisfying the first-order conditions:

$$du'(c^1) = v(\alpha^1), \tag{41}$$

$$du'(\hat{c}^1) = kv(\alpha^1). \tag{42}$$

Substituting for λ from (39) into (40), employing conditions (41) and (42), yields:

$$\frac{\pi^1 d v'(\alpha^1)}{v(\alpha^1)} \cdot \frac{1 - \delta\mu k}{1 - \mu k} = \frac{\pi^1 + \alpha^1 \frac{\partial \pi^1}{\partial \alpha^1}}{1 + \alpha^1 \frac{\partial \pi^1}{\partial y^1}}, \tag{43}$$

where $\delta \equiv \hat{\pi}^2 / \pi^1 > 1$ which follows from the first-order conditions in (41) and (42) by virtue of the strict concavity of u and as $k > 1$. Since $\delta > 1$ it further follows that $1 - \mu k > 1 - \delta\mu k > 0$, hence $\frac{1-\delta\mu k}{1-\mu k} < 1$. Thus, comparing (43) with (36) demonstrates the downward distortion on α^1 due to the binding incentive compatibility constraint.[23]

The equilibrium labor market contract offered to type 1 workers in the asymmetric information case is denoted by (y^{1*}, α^{1*}) and is determined by the solution to the system of equations

[23] To see this formally, notice that (43) can be written as $\frac{\pi^1 dv'(\alpha^1)}{v(\alpha^1)} \cdot \frac{1-\delta\mu k}{1-\mu k} - \frac{\pi^1 + \alpha^1 \frac{\partial \pi^1}{\partial \alpha^1}}{1+\alpha^1 \frac{\partial \pi^1}{\partial y^1}} = 0$. This can be interpreted as a first order condition along the envelope (when all else has been chosen optimally by the firm and by the workers). Then it follows that $\frac{\pi^1 dv'(\alpha^1)}{v(\alpha^1)} > \frac{\pi^1 + \alpha^1 \frac{\partial \pi^1}{\partial \alpha^1}}{1+\alpha^1 \frac{\partial \pi^1}{\partial y^1}}$ since $\frac{1-\delta\mu k}{1-\mu k} < 1$. Comparing this with (36) (in which the two terms are equalized) assuming that the second order conditions are satisfied, implies that the α^1 under the asymmetric information case (for which we obtain an inequality) is lower than the α^1 under the symmetric information case (for which we obtain an equality). This is because the derivative is positive at α^1 in the asymmetric information allocation and hence the optimal choice of α under the symmetric information case should be higher (in order to obtain an equality, i.e. the first order condition).

(39)–(42). Moreover, we denote by $\pi^{1*} = \pi^1(y^{1*}, \alpha^{1*})$ and $\hat{\pi}^{2*} = \pi^2(y^{1*}, \alpha^{1*})$ the associated optimal (expected) number of children chosen by type 1 workers and type 2 mimickers in response to the equilibrium contract.

6.3.3 Equilibrium with Parental Leave Mandate

We turn next to examine the potentially efficiency-enhancing role played by imposing a binding parental leave rule. A binding parental leave rule sets a lower bound on the duration of parental leave offered by the firm. Denote by $\bar{\alpha}$, the lower bound set by the government for the duration of parental leave, where $\alpha^{1*} < \bar{\alpha} < \alpha^{2*}$, and $\alpha^{i*}, i = 1, 2$, denotes the duration of parental leave offered to type-i workers in a separating equilibrium under asymmetric information. Notice that the duration could potentially be set to $\bar{\alpha} \geq \alpha^{2*}$, which would implement a pooling (rather than a separating) allocation, a scenario which is not possible under the laissez–faire regime. However, by assumption, type 1 workers strictly prefer their separating equilibrium bundle to any pooling allocation (this assumption guarantees the existence of the separating equilibrium under the laissez–faire regime). A pooling allocation, therefore, can never attain a Pareto improvement relative to the separating allocation under the laissez–faire regime. Thus, our assumption that $\bar{\alpha} < \alpha^{2*}$ is without loss of generality. Notice further that, similar to the quasi-linear specification, setting a binding parental leave rule implies that firms hiring type 1 workers will derive positive profits in equilibrium. The latter will sustain in equilibrium, despite the threat of entry, due to the binding parental leave rule. As in the quasi-linear specification, we assume that the government levies a confiscatory profit tax on firms' profits and further assume that the tax revenues are rebated across the board in a lump-sum fashion.

The labor contract offered to type 1 workers in a separating equilibrium, in the presence of a binding parental leave rule supplemented by a confiscatory profit tax and a universal lump-sum transfer, will be given by the solution to the following constrained optimization program:

$$L = \max_{y^1,\lambda,\mu} \left[\max_{\pi^1} [u(y^1 - d\pi^1 + T) + \pi^1 v(\bar{\alpha})] + \lambda(1 - \pi^1\bar{\alpha} - y^1) + \mu\left(U^{2*}(T) - \max_{\hat{\pi}^2} [u(y^1 - d\hat{\pi}^2 + T) + \hat{\pi}^2 k v(\bar{\alpha})] \right) \right], \quad (44)$$

where λ and μ denote the Lagrange multipliers associated, correspondingly, with the type 1 zero profit condition and the type 2 incentive compatibility constraint, T denotes the universal lump sum transfer and $U^{2*}(T)$ denotes the utility level associated with the labor contract offered to type 2 workers in equilibrium, formally given by the solution to the following constrained maximization program:

$$U^{2*}(T) = \max_{y^2,\alpha^2,\eta} \left[\max_{\pi^2} [u(y^2 - d\pi^2 + T) + \pi^2 k v(\alpha^2)] + \eta(1 - \pi^2\alpha^2 - y^2) \right], \quad (45)$$

where η denotes the Lagrange multiplier associated with the type 2 zero profit condition.

Notice that by virtue of the binding parental leave mandate, the type 1 zero profit constraint will be slack in the optimal solution for the maximization program in (44). Moreover, the type 2 incentive compatibility constraint will bind in the optimal solution for the maximization program.[24]

Denote by $U^1(\bar{\alpha}, T)$ the utility level associated with the labor contract offered to type 1 workers in equilibrium, equal to the Lagrangean expression L in (44). Consider now the following maximization program:

$$\max_{\bar{\alpha},T,\phi} \left[U^1(\bar{\alpha}, T) + \phi(1 - \pi^1\bar{\alpha} - y^1 - T/\gamma^1) \right], \tag{46}$$

where ϕ denotes the Lagrange multiplier associated with the government revenue constraint, which states that the total revenues raised by the confiscatory profit taxation on firms hiring type 1 workers, given by $\gamma^1(1 - \pi^1\bar{\alpha} - y^1)$, weakly exceeds the total amount of transfers, given by T, recalling that the total population size is normalized to one.

Notice that the parental leave mandate is not binding for the labor contract offered to type 2 workers. Further notice that type 2 workers receive a positive transfer, $T > 0$, from the government (financed by the confiscatory profit tax on firms employing type 1 workers). Thus, relative to the laissez–faire allocation, type 2 workers become unambiguously better-off. In order to attain a Pareto improvement, hence, one has to show that the binding parental leave mandate (supplemented by the confiscatory profit taxation and a uniform lump-sum transfer) makes type 1 workers (weakly) better-off. The following proposition states a necessary and sufficient condition for obtaining a Pareto improvement relative to the benchmark allocation in the extended model.

Proposition 3. *A Pareto improvement exists if-and-only-if*

$$\frac{\gamma^2}{\gamma^1} < \frac{\frac{u'(c^{2*})\left(1+\alpha^{1*}\frac{\partial \pi^1}{\partial y^1}\right)}{1+\alpha^{2*}\frac{\partial \pi^2}{\partial y^2}} \left[\frac{\pi^{1*}v'(\alpha^{1*})d}{v(\alpha^{1*})} - \frac{\pi^{1*}+\alpha^{1*}\frac{\partial \pi^1}{\partial \bar{\alpha}}}{1+\alpha^{1*}\frac{\partial \pi^1}{\partial y^1}} \right]}{kv'(\alpha^{1*}) \cdot (\hat{\pi}^{2*} - \pi^{1*})} \tag{47}$$

where (y^{1*}, α^{1*}) *and* (y^{2*}, α^{2*}) *are associated with the maximization problems* (44) *and* (45)*, respectively.*

Proof *See appendix E.* □

The inequality condition in (47) is a generalization of the necessary and sufficient condition stated in Proposition 1, allowing for nonlinear utility of consumption and endogenous π. The expression in brackets on the numerator of (47) measures the magnitude of the downward distortion on α^1, due to the binding incentive compatibility constraint rendering type 2 workers

[24]This is shown in appendix E.

just indifferent between mimicking their type 1 counterparts or not. It is positively signed by virtue of (43) and works in the direction of making it more likely to have a Pareto improvement. The expression in the denominator is also positively signed and captures the information rent associated with type 2 mimickers. This term works in the direction of making it less likely for a Pareto improvement to occur. Resolving the ambiguity depends on the parametric assumptions.

Before closing this section, we briefly comment on the possibility to use another policy instrument, namely, subsidies that decrease the costs of having children. Subsidizing, say, child care costs, would entail a reduction in the parameter d. This would be similar to the subsidized parental leave studied in section 4.3 with the exception that π is here endogenous. In similarity with that analysis, there would be a mechanical effect implying a tightening of the incentive constraint (because a child care subsidy would be more valued by type 2 workers for whom the incentive constraint is binding). In addition, there would be a behavioral effect changing the difference in π between the two agents. As we demonstrated in appendix B the effect of a change the difference in π on the possibility to attain a Pareto improvement is generally ambiguous (the distortion becomes larger, but the information rent that has to be provided to type 2 workers becomes larger as well). Hence, we cannot say whether a subsidy to child care would be beneficial on efficiency grounds. From an equity perspective, however, a subsidy to child care will always be desirable as long as the welfare weight attach to the utility of type 2 workers is sufficiently large (see section 5).

7 Conclusions

Despite a remarkable post-war convergence process, substantial gender differences in pay and employment levels are prevalent in most OECD countries. A major factor that contributes to the persistent gender gaps in labor market performance is women's traditional role in the household. Child-related absences from work imply that women tend to accumulate less job experience, are more prone to career discontinuity, and typically compromise on part-time flexible non-professional jobs, resulting in a substantial motherhood wage penalty. Women are essentially trading off flexibility for compensation in order to reconcile household and work obligations. Workplace flexibility is to a large extent shaped by government policy, with a notable example being parental leave mandates. In this paper we have employed a theoretical model capturing the gender-driven career/family segmentation of the labor market, and used it to present a novel normative justification for parental leave rules.

We have set focus on a competitive labor market in which firms cannot distinguish between workers who differ in their career/family orientation. This reflects either asymmetric information between workers and firms, or, an inability to tag based on observable attributes correlated with workers' (unobserved) career/family orientation, such as age, gender, marital status, number of children, and so on, due to anti-discrimination legislation. We have demonstrated how this can result in an under-provision of workplace flexibility and differences in wages between equally skilled men and women. In this setting, we have highlighted how parental leave arrangements can be a key policy tool to regulate the extent of workplace flexibility and serve a dual role of correcting for the market failure associated with the under-provision of workplace flexibility and promoting redistributive goals by reducing gender pay gaps.

References

Angelov, N., Johansson, P., and Lindahl, E. (2016). Parenthood and the gender gap in pay. *Journal of Labor Economics*, 34(3):545–579.

Balestrino, A., Cigno, A., and Pettini, A. (2002). Endogenous Fertility and the Design of Family Taxation. *International Tax and Public Finance*, 9(2):175–193.

Barigozzi, F., Cremer, H., and Roeder, K. (2017). Women's career choices, social norms and child care policies. IZA Discussion Paper No. 10502.

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

Bastani, S., Blumkin, T., and Micheletto, L. (2016). Anti-discrimination legislation and the efficiency-enhancing role of mandatory parental leave. CESifo Working Paper No. 5917.

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.

Blau, F. D. and Kahn, L. M. (2013). Female labor supply: Why is the US falling behind? *The American Economic Review*, 103(3):251–256.

Blundell, R., Costa Dias, M., Meghir, C., and Shaw, J. (2016). Female labor supply, human capital, and welfare reform. *Econometrica*, 84(5):1705–1753.

Bronson, M. A. (2015). Degrees are forever: Marriage, educational investment, and lifecycle labor decisions of men and women. mimeo, Georgetown University.

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

Chan, J. and Eyster, E. (2003). Does banning affirmative action lower college student quality? *American Economic Review*, 93(3):858–872.

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

Cremer, H. and Roeder, K. (2017). Social insurance with competitive insurance markets and risk misperception. *Journal of Public Economics*, 146:138–147.

Del Rey, E., Racionero, M., and Silva, J. I. (2017). On the effect of parental leave duration on unemployment and wages. *Economics Letters*, 158:14–17.

Eckstein, Z., Eichenbaum, M., and Peled, D. (1985). Uncertain lifetimes and the welfare enhancing properties of annuity markets and social security. *Journal of Public Economics*, 26(3):303 – 326.

Encinosa, W. (2001). A comment on Neudeck and Podczeck's 'adverse selection and regulation in health insurance markets'. *Journal of Health Economics*, 20(4):667–673.

Finkelstein, A. (2004). Minimum standards, insurance regulation and adverse selection: evidence from the medigap market. *Journal of Public Economics*, 88(12):2515–2547.

Fryer, R. G., Loury, G. C., and Yuret, T. (2008). An economic analysis of color-blind affirmative action. *Journal of Law, Economics, and Organization*, 24(2):319–355.

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

McFadden, D., Noton, C., and Olivella, P. (2015). Minimum coverage regulation in insurance markets. *SERIEs*, 6(3):247–278.

Neudeck, W. and Podczeck, K. (1996). Adverse selection and regulation in health insurance markets. *Journal of Health Economics*, 15(4):387 – 408.

Olivetti, C. and Petrongolo, B. (2016). The Evolution of Gender Gaps in Industrialized Countries. *Annual Review of Economics*, 8:405–434.

Pichler, S. and Ziebarth, N. R. (2017). The pros and cons of sick pay schemes: Testing for contagious presenteeism and noncontagious absenteeism behavior. *Journal of Public Economics*, 156:14 – 33.

Rossin-Slater, M. (2017). Maternity and Family Leave Policy. NBER Working Paper No. 23069.

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.

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

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

Thomas, M. (2018). The impact of mandated maternity benefits on the gender differential in promotions: Examining the role of adverse selection. mimeo, Cornell University.

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

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, $\hat{U}^1(\overline{\alpha}) = 1 - \pi^1 \overline{\alpha} - \frac{\gamma^2}{\gamma^1} \cdot T(\overline{\alpha}) + \pi^1 v(\overline{\alpha})$.

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 Comparative statics with respect to π

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

Recall that condition (9) was expressed in terms of the quantities characterizing the market equilibrium in the benchmark case. 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 his/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:

$$v'(\alpha^2) = 1, \tag{48}$$

$$1 - \alpha^2\pi^2 + \pi^2 v(\alpha^2) = 1 - \alpha^1\pi^1 + \pi^2 v(\alpha^1). \tag{49}$$

Now fix π^2 and consider (49). Since α^2 is given by the implicit solution to (48), the LHS of (49) expression does not depend on π^1. Total differentiation of (49) 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

$$\alpha_1 = \frac{\partial \alpha^1}{\partial \pi^1}\left[\pi^2 v'(\alpha^1) - \pi^1\right]. \tag{50}$$

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 he/she works more than the efficient amount) implies that:

$$\frac{\partial \alpha^1}{\partial \pi^1} > 0 \quad \text{and} \quad \frac{\partial c^1}{\partial \pi^1} < 0. \tag{51}$$

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

$$\gamma^2/\gamma^1 < \frac{\left[1 - \frac{1}{v'(\alpha^1)}\right]}{\frac{\pi^2}{\pi^1} - 1}. \tag{52}$$

It can immediately be seen that, for π^2 fixed, a decrease in π^1 implies that the denominator in (52) increases, which works in the direction of making it less likely for the government to achieve a Pareto improvement. Moreover, we know from (51) that a decrease in π^1 implies that α^1 decreases. Thus, the numerator $\left[1 - \frac{1}{v'(\alpha^1)}\right]$ in (52) 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 (52) 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 have expected that, as the difference between π^1 and π^2 becomes larger, the distortion that arises due to asymmetric information (or do the inability to use tagging due to anti-discrimination legislation) increases, and thus the scope for government intervention would become larger. This intuition is reflected in 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.

C Numerical example

In this section we provide a numerical example that illustrates the possibility to simultaneously satisfy the condition for Pareto-improvement in Proposition 1 and the existence condition for a separating equilibrium discussed on page 10. The numerical example also sheds light on the analytical ambiguitive of the comparative statics w.r.t. π discussed in appendix B.

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

Assuming $v(\alpha) = \frac{\alpha^\eta}{\eta}, \eta > 0$, the conditions defining the benchmark equilibrium take the form:

$$\alpha^{2(\eta-1)} = 1 \Longleftrightarrow \alpha^2 = 1 \tag{53}$$

$$u^2 = 1 - \pi^1\alpha^1 + \frac{\pi^2\alpha^{1\eta}}{\eta}, \quad \text{where} \quad u^2 = 1 - \pi^2\alpha^2 + \frac{\pi^2\alpha^{2\eta}}{\eta} = \frac{(1-\eta)\,\pi^2}{\eta} + 1 \tag{54}$$

$$c^1 = 1 - \pi^1\alpha^1, \tag{55}$$

$$c^2 = 1 - \pi^2\alpha^2 = 1 - \pi^2. \tag{56}$$

Notice that condition (53) determines the efficient amount of parental leave offered to type 2 workers; condition (54) is the binding (IC2) constraint which renders type 2 workers indifferent between mimicking type 1 or sticking to their contract, and conditions (55) and (56) state the consumption levels associated with type 1 and type 2 workers, respectively, determined by the corresponding zero profit conditions.

From (53)-(56) it can be derived that α^{1*} is given by the (unique) implicit solution to:

$$\frac{\alpha^1\eta}{(\alpha^{1\eta} - (1-\eta))} = \pi^2/\pi^1. \tag{57}$$

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

$$\gamma^2/\gamma^1 < \frac{1-(\alpha^{1*})^{\eta-1}}{\pi^2/\pi^1 - 1}. \tag{58}$$

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

The separating equilibrium exists only when the fraction of type 2 workers in the population exceeds a certain threshold.[25] This threshold ensures that type 1 workers strictly prefer the bundle intended for them in the benchmark separating market equilibrium to any bundle associated with a pooling allocation. 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:

$$\gamma^1 + \gamma^2 = 1, \tag{59}$$

$$\frac{1}{\pi^1\alpha^{(b-1)}} = 1/(\gamma^1\pi^1 + \gamma^2\pi^2), \tag{60}$$

$$1 - \left(\gamma^1\pi^1 + \gamma^2\pi^2\right)\alpha + \frac{\pi^1\alpha^b}{b} = 1 - \pi^1\alpha^1 + \frac{\pi^1{\alpha^1}^b}{b}, \tag{61}$$

where α^1 is the α for type 1 which prevails in the separating equilibrium and is given by the solution to (57). Denoting the solution to (59)-(61) by the triplet ($\hat{\gamma}^1$,$\hat{\gamma}^2$, $\hat{\alpha}$), a separating equilibrium exists if-and-only-if:

$$\gamma^2/\gamma^1 \geq \hat{\gamma}^2/\hat{\gamma}^1. \tag{62}$$

We now proceed to numerically analyze the possibility to simulatenously satisfy equations (58) and (62). For this purpose, we assume the utility from parental leave is CRRA, $v(\alpha) = \frac{\alpha^\eta}{\eta}$, where $0 < \eta < 1$ to ensure concavity.

In figure 5 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

[25] See Rothschild and Stiglitz (1976).

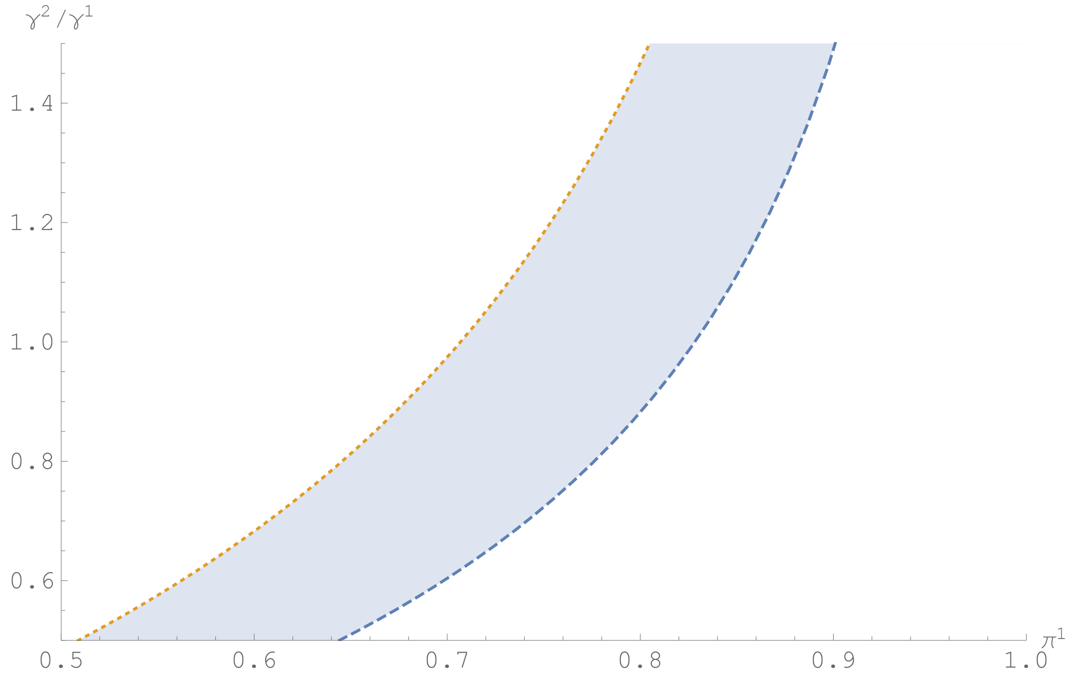


Figure 5: Numerical illustration of a region where the existence condition and the condition for Pareto-improvement are simultaneously satisfied.

wide range of parameter combinations.[26]

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.[27] 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 η 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, $1-\eta$.

D Proof of proposition 2

D.1 Two Lemmas

We begin by proving two lemmas that characterize the optimal duration of parental leave associated with a separating equilibrium and pooling equilibrium, respectively. We then prove

[26]Notice that according to our parametric specification, the necessary and sufficient condition (52) 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.

[27]To see this, consider equation (9) satisfied as an equality. The upward slope of the upper curve in figure 5 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.

proposition 2 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.[28]

Recall that the social maximization problem is defined as follows:

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

where $U^i(\overline{\alpha}, j)$ 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 4. To ease but slightly abuse notation, we will drop the second argument j of U in our exposition below, as it will always be obvious which equilibrium regime is under consideration.

Lemma 3 (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 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:

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

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

[28] This assumption is not necessary but is made for simplicity. We comment on how it affects the results in footnote 29.

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 $\beta^1 \equiv \beta$) :

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

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. □

Lemma 3 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) of Lemma 3, 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, a point will eventually be reached where an increase in 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.[29,30]

[29]As mentioned on page 40, in our derivations we assume that the necessary and sufficient condition for Pareto improvement is satisfied. Without this assumption the characterization in Lemma 3 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).

[30]Notice that we have, just as in section 4, 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 subsidized parental leave (see our discussion in section 4.3) would further enhance the government capacity to redistribute from type 1 to type 2 workers. We then anticipate that the government will increase the generosity of the subsidized parental leave system as the weight assigned to family-oriented workers increases (alongside extending the duration of the parental leave). When full weight is assigned to career-oriented workers, there will be nothing to gain from a subsidized 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.

The next lemma characterizes the pooling regime.

Lemma 4 (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 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 propensity 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. □

Lemma 4 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, as it serves to mitigate the 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.

D.2 Proof of the proposition

We next combine Lemma 3 and Lemma 4 to prove Proposition 2 and characterize the social optimum as a function of the weight assigned to type 1 (career-oriented) workers, β. We first prove part (ii), then part (i), and finally part (iii).

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:

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

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.

E Proof of Proposition 3

Notice that by virtue of the binding parental leave mandate, the type 1 zero profit constraint will be slack in the optimal solution for the maximization program in (44). We now show that the incentive constraint for type 2 workers in (44) is binding. To see this, suppose by negation that the constraint is slack. Differentiation of the Lagrange expression in (44) with respect to y^1, assuming that (by our presumption) both constraints are slack (hence $\lambda = \mu = 0$), employing the type 1 worker envelope condition, yields:

$$\frac{\partial L}{\partial y^1} = u'(y^1 - d\pi^1 + T) > 0. \tag{63}$$

We obtain the desired contradiction to the presumed optimality. It follows that the incentive constraint is necessarily binding in the optimum. Denote by $y^1(\bar{\alpha}, T)$ the optimal solution for the maximization in (44), given by the implicit solution to the binding incentive constraint:

$$U^{2*}(T) = \max_{\hat{\pi}^2}[u(y^1 - d\hat{\pi}^2 + T) + \hat{\pi}^2 kv(\bar{\alpha})]. \tag{64}$$

Denoting by $\hat{U}^2(\bar{\alpha}, T) = \max_{\hat{\pi}^2}[u(y^1(\bar{\alpha}, T) - d\hat{\pi}^2 + T) + \hat{\pi}^2 kv(\bar{\alpha})]$ the RHS of the above equation (the utility derived by a mimicking type 2 worker), it is straightforward to verify, employing the type 2 worker envelope condition, that:

$$\frac{\partial \hat{U}^2(\bar{\alpha}, T)}{\partial y^1} = u'(y^1 - d\hat{\pi}^2 + T) > 0. \tag{65}$$

Thus, as $U^{2*}(T)$ is independent of y^1, (64) uniquely determines $y^1(\bar{\alpha}, T)$. Fully differentiating (64) with respect to $\bar{\alpha}$ and T, employing the type 2 worker envelope conditions and the first order conditions for the maximization program in (45) , yields, upon re-arrangement:

$$1 + \frac{\partial y^1(\bar{\alpha}, T)}{\partial T} = \frac{u'(c^2)}{\left(1 + \alpha^2 \frac{\partial \pi^2}{\partial y^2}\right) u'(\hat{c}^1)} > 0, \tag{66}$$

where $c^2 = y^2 - d\pi^2 + T$ and $\hat{c}^1 = y^1 - d\hat{\pi}^2 + T$. The inequality sign follows by virtue of the maximization of type 2 worker with respect to π^2, which implies that $du'(y^2 - d\pi^2 + T) =$

$kv(\bar{\alpha})$, so that $\frac{\partial \pi^2}{\partial y^2} = 1/d > 0$; and,

$$\frac{\partial y^1(\bar{\alpha}, T)}{\partial \bar{\alpha}} = -\frac{\hat{\pi}^2 kv'(\bar{\alpha})}{u'(\hat{c}^1)} < 0. \tag{67}$$

We turn next to show that the revenue constraint in the maximization program given in (46) is binding in the optimum. To see this, suppose by negation that the constraint is slack (hence $\phi = 0$). Differentiating $U^1(\bar{\alpha}, T)$ with respect to T, employing the type 1 worker envelope condition, yields:

$$\frac{\partial U^1(\bar{\alpha}, T)}{\partial T} = u'(c^1)\left(1 + \frac{\partial y^1}{\partial T}\right) > 0, \tag{68}$$

where $c^1 = y^1 - d\pi^1 + T$, and where the inequality sign follows from (66) . Thus, we obtain the desired contradiction to the presumed optimality. We therefore conclude that the revenue constraint is binding in the optimum. For each duration of parental leave, $\bar{\alpha}$, the level of transfer, $T(\bar{\alpha})$, is given by the implicit solution to the binding revenue constraint, namely:

$$G(\bar{\alpha}, T) \equiv 1 - \pi^1\bar{\alpha} - y^1 - T/\gamma^1 = 0. \tag{69}$$

Differentiating G with respect to T, yields upon re-arrangement:

$$\frac{\partial G}{\partial T} = -\left(1 + \frac{\partial y^1}{\partial T}\right)\left(1 + \bar{\alpha}\frac{\partial \pi^1}{\partial y^1}\right) - \frac{\gamma^2}{\gamma^1} < 0, \tag{70}$$

where the inequality sign follows from (66) and by virtue of the maximization of type 1 worker with respect to π^1, which states that $du'(y^1 - d\pi^1 + T) = v(\bar{\alpha})$, so that $\frac{\partial \pi^1}{\partial y^1} = 1/d > 0$. We thus conclude that $T(\bar{\alpha})$ is uniquely determined. By construction, $T(\bar{\alpha}) \to 0$ as $\bar{\alpha} \to \alpha^{1*}$; namely, the extent of cross-subsidization from type 1 to type 2 workers converges to zero as we approach the laissez–faire equilibrium.

The maximization program in (46) can be re-formulated as:

$$\max_{\bar{\alpha}} H(\bar{\alpha}), \tag{71}$$

where $H(\bar{\alpha}) \equiv U^1[\bar{\alpha}, T(\bar{\alpha})]$ and $T(\bar{\alpha})$ is given by the implicit solution to (69).

Assuming that the second-order condition for the maximization in (71) is satisfied, a necessary and sufficient condition for a Pareto improvement is:

$$\left.\frac{\partial H}{\partial \bar{\alpha}}\right|_{\bar{\alpha}=\alpha^{1*}, T=0} = \left.\frac{\partial U^1}{\partial \bar{\alpha}}\right|_{\bar{\alpha}=\alpha^{1*}, T=0} + \left.\left[\frac{\partial U^1}{\partial T} \cdot \frac{\partial T}{\partial \bar{\alpha}}\right]\right|_{\bar{\alpha}=\alpha^{1*}, T=0} > 0. \tag{72}$$

That is, starting at the laissez–faire allocation, setting a slightly binding parental leave rule (supplemented by a confiscatory profit tax and a uniform lump-sum transfer) increases the utility of type 1 workers.

Differentiation of $U^1(\bar{\alpha}, T)$ with respect to $\bar{\alpha}$ and T, employing type 1 worker envelope condition and evaluating the expressions at the laissez–faire allocation levels, yields:

$$\left.\frac{\partial U^1}{\partial \bar{\alpha}}\right|_{\bar{\alpha}=\alpha^{1*},T=0} = u(c^{1*})\left.\frac{\partial y^1}{\partial \bar{\alpha}}\right|_{\bar{\alpha}=\alpha^{1*},T=0} + \pi^{1*}v'(\alpha^{1*}) \tag{73}$$

and

$$\left.\frac{\partial U^1}{\partial T}\right|_{\bar{\alpha}=\alpha^{1*},T=0} = u(c^{1*})\left(1 + \left.\frac{\partial y^1}{\partial T}\right|_{\bar{\alpha}=\alpha^{1*},T=0}\right), \tag{74}$$

where $c^{1*} = y^{1*} - d\pi^{1*}$.

Differentiating G, given in (69) , with respect to $\bar{\alpha}$, yields upon re-arrangement:

$$\frac{\partial G}{\partial \bar{\alpha}} = -\left[\pi^1 + \bar{\alpha}\frac{\partial \pi^1}{\partial \bar{\alpha}} + \frac{\partial y^1}{\partial \bar{\alpha}}\left(1 + \bar{\alpha}\frac{\partial \pi^1}{\partial y^1}\right)\right]. \tag{75}$$

Thus, by virtue of (70), it follows that:

$$\left.\frac{\partial T}{\partial \bar{\alpha}}\right|_{\bar{\alpha}=\alpha^{1*},T=0} = -\left.\frac{\frac{\partial G}{\partial \bar{\alpha}}}{\frac{\partial G}{\partial T}}\right|_{\bar{\alpha}=\alpha^{1*},T=0} = -\frac{\left[\pi^{1*} + \alpha^{1*}\left.\frac{\partial \pi^1}{\partial \bar{\alpha}}\right|_{\bar{\alpha}=\alpha^{1*},T=0} + \left.\frac{\partial y^1}{\partial \bar{\alpha}}\right|_{\bar{\alpha}=\alpha^{1*},T=0}\left(1 + \alpha^{1*}\left.\frac{\partial \pi^1}{\partial y^1}\right|_{\bar{\alpha}=\alpha^{1*},T=0}\right)\right]}{\left(1 + \left.\frac{\partial y^1}{\partial T}\right|_{\bar{\alpha}=\alpha^{1*},T=0}\right)\left(1 + \alpha^{1*}\left.\frac{\partial \pi^1}{\partial y^1}\right|_{\bar{\alpha}=\alpha^{1*},T=0}\right) - \frac{\gamma^2}{\gamma^1}}. \tag{76}$$

Substitution for $\left.\frac{\partial U^1}{\partial \bar{\alpha}}\right|_{\bar{\alpha}=\alpha^{1*},T=0}$, $\left.\frac{\partial U^1}{\partial T}\right|_{\bar{\alpha}=\alpha^{1*},T=0}$ and $\left.\frac{\partial T}{\partial \bar{\alpha}}\right|_{\bar{\alpha}=\alpha^{1*},T=0}$ from (73) , (74) and (76) into (72) employing (41) , (42) , (66) and (67), yields upon re-arrangement (we abbreviate notation by referring to all derivatives as evaluated at the laissez–faire levels):

$$\left.\frac{\partial H}{\partial \bar{\alpha}}\right|_{\bar{\alpha}=\alpha^{1*},T=0} > 0 \Longleftrightarrow \frac{\frac{u'(c^{2*})\left(1+\alpha^{1*}\frac{\partial \pi^1}{\partial y^1}\right)}{1+\alpha^{2*}\frac{\partial \pi^2}{\partial y^2}}\left[\frac{\pi^{1*}v'(\alpha^{1*})d}{v(\alpha^{1*})} - \frac{\pi^{1*}+\alpha^{1*}\frac{\partial \pi^1}{\partial \bar{\alpha}}}{1+\alpha^{1*}\frac{\partial \pi^1}{\partial y^1}}\right]}{kv'(\alpha^{1*})\cdot(\hat{\pi}^{2*} - \pi^{1*})} > \frac{\gamma^2}{\gamma^1}. \tag{77}$$

F The role of subsidized leave

Suppose that the government provides a flat subsidized parental leave system, which takes the form: $G^i = b\pi^i$, $i = 1, 2$, where $b > 0$. We will characterize the modified necessary and sufficient condition for attaining a Pareto improvement. We then demonstrate that the set of parameters for which a Pareto improvement is attained is a subset of the corresponding set for the regime with a universal transfer. Thus, providing a subsidized scheme which depends on the time spent on leave, rather than a universal system, results in a shrinkage of the set of parameters for which a Pareto improvement can be obtained.

A separating equilibrium allocation associated with a mandatory parental leave rule $\overline{\alpha}$, $\alpha^{1*} \leq \overline{\alpha} \leq \alpha^{2*}$, supplemented by a confiscatory tax levied on pure profits, T, and a flat subsidized transfer, b, is given by: $\{\alpha^i, y^i\}_{i=1,2}, T, b$,

where:

$$y^i = 1 - \pi^i\alpha^i; i = 1, 2, \tag{78}$$
$$\alpha^1 = \overline{\alpha}, \tag{79}$$
$$\alpha^2 = \alpha^{2*} \quad \text{where} \quad v'(\alpha^{2*}) = 1, \tag{80}$$
$$y^2 + b\pi^2 + \pi^2 v\left(\alpha^2\right) = y^1 - T + b\pi^2 + \pi^2 v\left(\alpha^1\right), \tag{81}$$
$$\gamma^1 T = b \cdot (\gamma^1\pi^1 + \gamma^2\pi^2). \tag{82}$$

Substituting for α^i and y^i, $i = 1, 2$, from conditions (78)-(80) into (81) and re-arranging, yields: $T(\overline{\alpha}) = \pi^2\alpha^{2*} - \pi^1\overline{\alpha} + \pi^2[v(\overline{\alpha}) - v(\alpha^{2*})]$. Notice that $T(\alpha^{1*}) = 0$ by construction of the laissez-faire 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 $\overline{\alpha} > \alpha^{1*}$.

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, $\hat{U}^1(\overline{\alpha}) = 1 - \pi^1\overline{\alpha} - T(\overline{\alpha}) + b\pi^1 + \pi^1 v(\overline{\alpha})$. Substituting for $T(\overline{\alpha})$ into (v) yields: $b(\overline{\alpha}) = \gamma^1 T(\overline{\alpha})/(\gamma^1\pi^1 + \gamma^2\pi^2)$. Substituting for $T(\overline{\alpha})$ and $b(\overline{\alpha})$ into $\hat{U}^1(\overline{\alpha})$ yields upon re-arrangement:

$$\hat{U}^1(\overline{\alpha}) = 1 - \pi^1\overline{\alpha} + \pi^1 v(\overline{\alpha}) - \left(\pi^2\alpha^{2*} - \pi^1\overline{\alpha} + \pi^2[v(\overline{\alpha}) - v(\alpha^{2*})]\right) \cdot \frac{\gamma^2\pi^2}{\gamma^1\pi^1 + \gamma^2\pi^2}.$$

By a similar reasoning to the one provided in the proof of proposition 1, a necessary and sufficient condition for a Pareto improvement under the flat subsidized system is $\left.\frac{\partial\hat{U}^1(\overline{\alpha}))}{\partial\overline{\alpha}}\right|_{\overline{\alpha}=\alpha^{1*}} > 0$. Differentiation of $\hat{U}^1(\overline{\alpha})$ with respect to $\overline{\alpha}$, evaluating the derivative at $\overline{\alpha} = \alpha^{1*}$, yields upon rearrangement:

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

The necessary and sufficient condition for the universal transfer regime, stated in proposition 1, is given by:

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

Comparing the two inequality conditions, recalling that $\pi^2 > \pi^1$, implies that the set of parameters for which a Pareto improvement is attained under the flat subsidized regime is indeed a subset of the set of parameters for which a Pareto improvement is attained under the universal transfer regime. This establishes our argument.