ISSN 2282-6483

Alma Mater Studiorum - Università di Bologna
DEPARTMENT OF ECONOMICS

Competition Between For-Profit and Non-Profit Firms: Incentives, Workers' Self-Selection, and Wage Differentials

Francesca Barigozzi
Nadia Burani

Quaderni - Working Paper DSE N°1072

Competition Between For-Profit and Non-Profit Firms: Incentives, Workers' Self-Selection, and Wage Differentials

Francesca Barigozzi* and Nadia Burani†

University of Bologna

July 2016

Abstract

We study optimal non-linear contracts offered by two firms competing for the exclusive services of workers, who are privately informed about their ability and motivation. Firms differ in their organizational form, and motivated workers are keen to be hired by the non-profit firm because they adhere to its mission. If the for-profit firm has a competitive advantage over the non-profit firm, the latter attracts fewer high-ability workers with respect to the former. Moreover, workers exert more effort at the for-profit than at the non-profit firm despite the latter distorts effort levels upwards. Finally, a wage penalty emerges for non-profit workers which is partly due to compensating effects (labor donations by motivated workers) and partly due to the negative selection of ability into the non-profit firm. The opposite results hold when it is the non-profit firm that has a competitive advantage.

JEL classification: D82, D86, J24, J31, M55.

Key-words: non-profit firms, multi-principals, intrinsic motivation, skills, bidimensional adverse selection, wage differential.

1 Introduction

According to data from the 2014 CEO Compensation Study, reported by Charity Navigator,[1] there are top executives of U.S. non-profit organizations whose annual compensation exceeds one million dollar.[2]

*Department of Economics, University of Bologna, P.zza Scaravilli 2, 40126 Bologna (Italy). E-mail: francesca.barigozzi@unibo.it

†Department of Economics, University of Bologna, Strada Maggiore 45, 40125 Bologna (Italy). E-mail: nadia.burani@unibo.it. Tel: +39 0512092642. Corresponding author.

[1] See Charity Navigator "2014 Charity CEO Compensation Study" at www.charitynavigator.org

[2] Among them, one can find the President of Chicago University and the President and CEO of Shedd Aquarium in Chicago or the President of the Icahn School of Medicine at Mount Sinai (N.Y.C.) or else the President of the Lincoln

Considering the comments to these data that one can find on the media, two different perspectives emerge. On the one hand, there is wide consensus by the public that seven-figure salaries are excessive for employees of non-profit organizations that receive private donations and, possibly, public funding. That's why, in various federal states such as New Jersey, New York, Florida and Massachusetts, there have been recent proposals to introduce legislation that would cap top managers' compensations at non-profits.[3] Indeed, "if you are going to work at a non-profit, you should have as your primary motivation the public good (...) and you would accept less compensation...".[4] On the other hand, some contributors claim that "just because someone works for a non-profit, doesn't necessarily mean they're doing it for free".[5] First it is pointed out that most non-profits tend to pay less than for-profit businesses for similar competencies. Moreover, it is highlighted that the non-profit organizations that pay the highest compensations are multi-million dollar operations: leading one of them requires individuals that possess extensive management expertise together with a thorough understanding of the issues that are unique to the non-profit's mission. Therefore, "attracting and retaining that type of talent requires a competitive level of compensation as dictated by the marketplace".[6]

Our analysis contributes to this debate, challenging the idea that non-profit employees, especially at the top of the wage ladder, should be ready to accept low salaries. In particular, we show that the competition between for-profit and non-profit organizations to attract the most talented workers, without *a priori* knowing their skills, tends to drive all salaries up. We also analyse how this interacts with the workers' willingness to donate part of their labor to non-profit organizations whose mission or goal they adhere to.

There exists a well-established empirical evidence on compensating wage differentials asserting that differences in wages across sectors or jobs are generated by differences in job characteristics or attributes for which heterogeneous workers have different willingnesses to pay. For instance, an earnings penalty attributed to compensating factors has been documented for public firms as opposed to private ones and for not-for-profit firms relative to for-profit organizations.[7]

Intrinsic motivation for being employed by non-profit or mission-oriented firms has often been viewed as a possible source of compensating wage differentials. This idea has been first proposed by Handy and Katz (1998) for non-profit vs for-profit managers, by Heyes (2005) in the health sector, and by Delfgaauw

Center for Performing Arts.

[3] These bills have not reached legislative approval yet.

[4] See: "Is A One Million Dollar Nonprofit CEO Salary As Bad As It Sounds", Forbes, January 23, 2013.

[5] See The Street: "These 9 Nonprofit Executives Made Over 1-Million Dollar", September 15, 2014.

[6] See the "2014 CEO Compensation Study" by Charity Navigator.

[7] For compensating wage differentials see Rosen (1986). The case of public versus private firms has been studied by Disney and Gosling (1998) and Melly (2005), among others. Lower average wages in not-for-profit firms relative to for-profit ones have been found by Preston (1989), Gregg *et al.* (2011).

and Dur (2007). Besley and Gathak (2005) consider in particular mission-oriented firms that operate in specific sectors (education, health and defence) and produce collective goods. Bénabou and Tirole (2010) highlight the role of firms' corporate social responsibility: some firms take employee-friendly or environment-friendly actions, some employers are mindful of ethics, or they even have an investor-friendly behavior (as ethical banks). All those organizations have in common the pursuit of a mission or goal that is valuable for some workers, precisely those who share such objectives and who are characterized by non-pecuniary motivations, together with the standard extrinsic incentives. The theoretical prediction is that relatively low pay and weak monetary incentives endogenously emerge in jobs where intrinsic motivation matters. Empirically, the so-called labor donative hypothesis as the determinant of compensating wage differentials (see Preston 1989) has been tested by Leete (2001) and Jones (2015), among others.

However, another strand of empirical work points out that the wage differential might arise because of a selection bias, given that a wage gap can also reflect *unobservable* differences in workers' ability across sectors or firms. See, for instance, Goddeeris (1988) for lawyers, Roomkin and Weisbrod (1999) for hospitals, and also Hwang *et al.* (1992) and Gibbons and Katz (1992).

Therefore, an open question still remains. Suppose that a wage penalty for workers employed in mission-oriented and non-profit sectors or firms is measured, although neither workers' intrinsic motivation nor ability can be directly observed: then, wages can be lower either because of the lower reservation wages of motivated workers or because of the lower productivity of workers self-selecting into such sectors or firms (or because of a combination of these two effects). In other words, when workers' productivity and motivation are the workers' private information, is it possible to disentangle the pure compensating wage differential from the selection effect of ability?

To this respect, we extend the analysis of Delfgaauw and Dur (2010) that studies the sorting of workers, who are heterogeneous in both their productivity and their public service motivation between the public and the private sector. This paper assumes that workers' characteristics are fully observable and that both sectors are fully competitive. We rather reckon that asymmetric information about potential applicants' traits, coupled with strategic interaction among employers, that are willing to attract the most talented and motivated workers, are key ingredients to model the situation we are interested in and to answer our research question.

In our paper, we consider a labor market characterized by two firms, a mission-oriented or non-profit firm and a standard for-profit firm.[8] The two firms compete to attract workers who are heterogeneous with respect to both their skills and their intrinsic motivation. These two characteristics are the workers' private information and are not correlated. In particular, workers can have either high or low ability,

[8] For expositional clarity, in the paper we consider the dichotomy non-profit vs for-profit but we consider the notions of non-profit and mission-oriented as almost equivalent. With a slight abuse of termilogy, our model could also be well-suited to study public vs private sector jobs and the wage differentials therein.

whereas motivation is continuously distributed in the unit interval. In order to elicit the applicants' private information, the two firms simultaneously offer screening contracts consisting in a non-linear wage which depends on the observable effort (task) level. Because of the strategic interaction between the two firms, the workers' outside options are type-dependent and endogenous and thus the analysis of a multi-principal framework with bidimensional screening is called for.

All workers experience a cost from effort provision, which can differ across workers types but which does not depend on the employer's organizational form. Conversely, motivated workers care about the mission pursued by the firm which employs them. More precisely, the payoff of motivated agents depends on their own type but also on the organizational form of the firm hiring them. When motivated workers are employed by the non-profit employer, they enjoy a non-monetary benefit (which is unrelated to effort exertion or output produced) because they share their organization's mission or goal. Therefore, the non-profit benefits from being able to attract motivated applicants, whose reservation wage is low, but this comes at the cost of having to sacrifice some payoff to engage in socially worthwhile projects.

The two firms have different objective functions, because the for-profit firm strictly maximizes its profits whereas the non-profit firm is assumed to face some constraints and is able to appropriate only a fraction of its revenues. Moreover, the two firms also differ in their technologies, which are characterized by different marginal products of labor, and in the prices they face on the final product market. For simplicity, we bunch all these sources of firms' heterogeneity into one single summary statistic, which is the firm's marginal revenue and we say that one firm has a competitive advantage over the other when its marginal revenue is higher than the rival's. We allow the non-profit organization to have a competitive advantage over the for-profit firm, despite its revenue constraints.

We are interested in the intensive rather than in the extensive margin: we focus on situations in which, in equilibrium, both firms are active and are able to attract a positive share of workers of each ability level.[9] We characterize the optimal incentive schemes offered by each firm, the sorting pattern of workers into firms, and relate it not only to the sign but also to the composition of the wage differential (disentangling the labor donative form the selection effect).

The optimal incentive contracts are based on workers' ability, whereas workers' motivation determines the labor supply from applicants of each ability level facing each firm. Optimal contracts differ according to how the difference in firm's marginal revenues relates to the difference in workers' skills. When the difference in marginal revenues is low compared to the heterogeneity in workers' skills the following happens: on the one hand, firms are very similar to each other and competition between them is fierce; on the other hand, skills are "distant" and this discourages mimicking between types with different abilities. So, optimal allocations (effort levels) are the efficient ones. Conversely, when the difference in marginal

[9] Instead, we disregard the instances in which only one firm is able to hire all workers of given skills, because these cases do not allow to examine the wage differential between firms.

revenues is high relative to the difference in workers' skills, the opposite occurs: competition between firms is less relevant and types are sufficiently close to each other so that mimicking becomes attractive. Therefore, internal incentive compatibility is the driving force in shaping optimal contracts and optimal allocations are distorted away from the first-best; in particular, the firm with a competitive disadvantage distorts high-ability workers' effort upwards and, eventually, the low-ability workers' effort is distorted downwards by the advantaged firm.

Moreover, the difference in firm's marginal revenues also determines the selection pattern of workers to firms. We show that the selection effect of ability is more pronounced under asymmetric information about workers' ability than when ability is perfectly observable. In particular, the higher is the difference in firm's marginal revenue, the higher is the share of high-ability workers and the lower is the share of low-ability workers accepting employment at the firm with a competitive advantage. For instance, when the non-profit organization has a competitive disadvantage, there exists negative selection of ability for the non-profit firm, which increases with the asymmetry between the two firms and which is exacerbated when incentive schemes are in place.

As for the wages offered by the two firms, our model is sufficiently rich to accommodate for both wage penalties and wage premia at the non-profit organization. We find that, when the non-profit firm has a competitive disadvantage, then a wage differential emerges in that the total salary gained by non-profit workers is lower than the salary that the same workers would gain if employed by the for-profit firm. Such a wage penalty for non-profit workers is always associated with lower effort provision. The result that workers' average ability is different across firms allows us to conclude that the earnings penalty possibly experienced by workers in non-profits is due in part to a true compensating wage differential (the labor donative hypothesis) and is in part driven by negative selection with respect to ability.[10] But our model also predicts that a non-profit wage premium can arise when the non-profit organization has a competitive advantage and when the positive selection effect of ability is sufficiently strong as to offset the labor donative effect.[11]

The rest of the paper is organized as follows. In the following subsection we describe the related literature. In Section 2, we set up the model; in Section 3, as benchmark case, we present the equilibrium when firms are perfectly informed about workers' ability. Section 4 introduces asymmetric information about ability and describes the equilibrium screening strategies of the two firms; the optimal sorting of workers into firms is considered together with the full characterization of the optimal contracts. Section 5 focuses on wage differentials and, finally, Section 6 concludes.

[10] This fact is consistent with the empirical evidence on the public-private wage gap documented in Bargain and Melly (2008) and with the for-profit vs non-profit wage differences found by Roomkin and Weisbrod (1999), among others.

[11] This finding confirms the empirical results found by Preston (1988), Borjas *et al.* (1983) and James (2002), among others.

1.1 Related literature

Our work contributes to two different strands of literature: from an economic point of view, it adds to the recent and rapidly growing literature on the self-selection of workers with intrinsic motivation into different firms/sectors of the labor market; from a theoretical point of view, it explicitly solves a multi-principal game in a labor market where two firms compete to attract workers who are characterized by two different dimensions of private information.

The problem of the design of optimal incentive schemes for intrinsically motivated workers has been tackled by Murdock (2002), Besley and Gathak (2005), Francois (2000 and 2003), Prendergast (2007) and Ghatak and Mueller (2011), whose attention has primarily been devoted to moral hazard and free-riding, while we consider the screening problem. Heyes (2005) and Delfgaauw and Dur (2007) are the first papers that address the issue of the selection of workers who are heterogeneous with respect to their motivation.

Previous results from theoretical literature admitting for workers' private information are ambiguous on whether mission-oriented firms or sectors are able to hire workers with lower or higher productivity. In particular, Handy and Katz (1998) find that lower wages attract managers that are more committed to the cause of the non-profit firm, but this comes at the cost of selecting less able managers.[12] Delfgaauw and Dur (2008) study the problem of workers' self-selection into public vs private sectors when the governmental agency designs screening contracts. Their screening mechanism is simplified because the public agency is constrained to hire at most two types of agents. They find that the public agency optimally hires the more dedicated and the laziest workers in the economy, but they are able to compare neither the workers' ability nor the wages across sectors.

The most closely related paper to ours is Delfgaauw and Dur (2010) that is framed in a full information setup. They show that the return to managerial ability is lower in the public than in the private sector. Hence, a public-private earnings differential exists, which is caused partly by a compensating wage differential (motivated workers evaluate more being employed in the public sector) and partly by selection arising endogenously (on average more productive workers enter the private sector where remuneration is higher). Our model extends the setup in Delfgaauw and Dur (2010) in two ways: first, bidimensional asymmetric information is considered rather than full information about the workers' characteristics and, second, firms interact strategically. Our model too accounts for the result of negative selection of workers' ability for the non-profit organization, coupled with the existence of a wage penalty for non-profit workers. On top of that, we also document either ability-neutrality or even positive selection of ability for the non-profit firm; the latter allows non-profit employees to enjoy a wage premium.

More recently, DeVaro et al. (2015) consider a non-profit firm, that is bound to offer flat wages,

[12] A limit of the analysis is that an exogenously given ranking is imposed for the levels of effort and for the reservation wages of different types of managers.

competing with perfectly competitive for-profit rivals in hiring workers. They show that workers hired by the non-profit firm have sufficiently high intrinsic motivation and that a wage differential favoring for-profit workers emerges when for-profit firms are more effective in training workers.[13] The matching of workers to firms is also analyzed by Kosfeld and von Siemens (2009, 2011) that model a competitive labor market with team production and adverse selection, where selfish and conditionally cooperative workers coexist. They show that workers separate in equilibrium, thereby leading to the emergence of heterogeneous "corporate cultures", like for-profit and non-profit. In addition, Auriol and Brilon (2014) consider two types of intrinsically motivated workers, good and bad workers, and show that non-profit organizations have to resort to higher monitoring to deter entry of bad workers, while for-profits increase both monitoring and bonus payments for pro-social behavior to contrast bad workers. Finally, Bénabou and Tirole (2016) study firms competing to attract workers who are heterogeneous with respect to their productivity and their work ethics. In a framework with multitasking and moral hazard, they show how competition for the most productive workers interacts with the incentive structure inside firms to undermine work ethics. This paper differs significantly from ours because it assumes an affine (rather than a non-linear) compensation scheme and it considers screening with respect to one dimension at a time (either productivity or work ethics).

From a technical point of view, our paper draws both from the literature on multidimensional screening and from the literature on multi-principals. Models where both problems are simultaneously considered are very few.

Screening when agents have several unobservable characteristics has been analyzed by some important papers that deal with continuous distributions of types: Armstrong (1996), Rochet and Choné (1998) and Basov (2005). They all show that it is almost impossible to extend to the multidimensional environment the qualitative results and the regularity conditions of the unidimensional case. Barigozzi and Burani (2016a) use a discrete two-by-two setup to study the bidimensional screening problem of a mission-oriented monopsonist willing to hire workers of unknown ability and motivation, when motivation is output-dependent.

The multi-principal literature with asymmetric information was initiated by the seminal contributions of Martimort (1992) and Stole (1992). Within this literature, the paper that is most closely related to ours is Rochet and Stole (2002) which extends the analysis carried out in Stole (1995) and studies duopolists competing in nonlinear prices in the presence of both vertical and horizontal preference uncertainty. Consumers are heterogeneous and privately informed about their preference for quality and about their outside opportunity cost. Contracts consist of quality-price pairs that only depend on consumers' (unidimensional) preference for quality. The outside opportunity cost, affects the consumers' decision

[13] The paper also tests the theoretical results with data on California establishments showing that for-profits firms offer higher wages and higher incentive pay with respect to non-profits.

about which firm to buy from.[14] We depart from Rochet and Stole (2002) because they only consider symmetric firms and find that incentive compatibility constraints are never binding for any firm, so that efficient quality allocation with cost-plus-fee pricing emerge as the equilibrium outcome.[15] Barigozzi and Burani (2016b) extends the framework of Barigozzi and Burani (2016a) to allow for competition between a non-profit and a for-profit hospital. There are two main differences between Barigozzi and Burani (2016b) and the present paper: first, the former considers a two-by-two setup in which workers' motivation can only take two possible values, then, most importantly, it considers output-oriented motivation, whereby optimal contracts depend on both workers' ability and motivation, not just on ability as in the present context.

Another closely related model is Biglaiser and Mezzetti (2000), which studies an incentive auction in which multiple principals bid for the exclusive services of an agent, who has private information about ability. It is shown that only downward incentive constraints, if any, might be binding and that the presence of multiple principals reduces the distortions in the agent's effort level. As opposed to Biglaiser and Mezzetti (2000), we show that the upward incentive constraint might also be binding and this leads to an upward distortion in the optimal allocation for high-ability workers.

2 The model

We consider a multi-principal setting with bidimensional asymmetric information. Two principals (firms) compete to hire agents (workers). Each agent (she) can work exclusively for one principal. Principals and agents are risk neutral.

Firms

Firms differ in the mission they pursue. One firm is profit-oriented while the other firm is a non-profit institution. Effort supplied by the agent is the only input the two firms need in order to produce output. We call x the *observable and measurable* effort (task) level that the agent is asked to provide.[16] Both firms' production functions display constant returns to effort so that the amount of output produced is

$$q_i\left(x\right) = c_i x$$

for each firm $i = F, N$, with F and N referring to the for-profit and to the non-profit firm, respectively.

[14] A similar setup is analyzed in Lehmann et al. (2014) that considers optimal nonlinear income taxes levied by two competing governments on individuals who are privately informed about their earnings capabilities (i.e. their skills) and their migration costs.

[15] Precisely the same result can be found in Armstrong and Vickers (2001) that model firms as supplying utility directly to consumers.

[16] In particular, x can be interpreted as a job-specific requirement like the amount of hours of labor or the amount of services the agent is asked to provide.

The marginal product of labor c_i is firm-specific, and we allow for the for-profit to be more efficient than the non-profit firm or vice-versa (see the concluding section for a discussion on empirical evidence).

Profit margins (per-worker, conditional on the worker being hired) are given by

$$\pi_i(x) = p_i \alpha_i q_i(x) - w_i(x) = p_i \alpha_i c_i x - w_i(x), \tag{1}$$

where $w_i(x)$ is the total wage or salary paid by firm i to the worker exerting effort x and where the price of output p_i is assumed to be exogenous. Again, we do not impose any exogenous ranking of output prices for the two firms, so that $p_F \gtreqless p_N$ (see below).

Importantly, parameter α_i captures the difference in revenue appropriation between the two firms.[17] We assume that $\alpha_F = 1$ whereas $0 < \alpha_N < 1$, implying that the non-distribution constraint which the non-profit firm is committed to limits the entrepreneur's ability to appropriate the firm's revenues. As we explain below, workers' motivational premium (which is enjoyed when they are hired by the non-profit firm) is precisely generated by the non-profit commitment to this revenue constraint. For instance, consider non-profit hospitals whose mission consists in providing care to both insured and uninsured patients. Then, a non-profit hospital is rewarded only for the fraction α_N of insured patients that it treats (compensated care), while its revenue is zero when treating uninsured patients (uncompensated/charity care).[18] Alternatively, our model could be rephrased in terms of a standard *vs* a mission-oriented firm, where the latter sacrifices some of its revenues or profits in the social interest: consider, for instance, socially responsible organizations (see Bénabou and Tirole 2010 and Kitzmueller and Shimshack 2012), or else ethical banks investing in social projects with low returns.

To economize on notation, let us set $k_i = p_i \alpha_i c_i$, so that $k_N = k_F$ describes a situation in which firms are symmetric with respect to technology and profit margins, whereas $k_i > k_j$ describes a situation in which firm i has a competitive advantage with respect to firm j.

The case in which the for-profit has a competitive advantage over the non-profit firm (i.e. $k_F > k_N$) is perhaps the most likely to occur. It arises when the two firms differ uniquely in their revenue appropriation, that implies $\alpha_N < \alpha_F = 1$, but are otherwise identical because they share the same technology and face the same output prices. This occurs when hospitals are paid a fixed tariff for every patient admitted for treatment, as in Diagnosis Related Group (DRG) systems, like Medicare in the U.S. or prospective payment systems in many European countries. It might also occur that the competitive advantage of the for-profit organization is even stronger because marginal products of labor are such that $c_N < c_F$, as when the non-profit firm is committed to employ, as a fraction of its workforce, people with disabilities or disadvantages and provide them with training and supportive services.

[17] A similar formulation of the objective function for non-profit firms was proposed in Glaeser and Shleifer (2001), following the ideas expressed in Hansmann (1996). In Glaeser and Shleifer (2001), nonetheless, non-profit firms have *profit* -rather than revenue- constrains.

[18] Think also about non-profit universities that provide total waivers of tuition fees to poor and promising students.

But it might also be that the non-profit has a competitive advantage over the for-profit firm (i.e. $k_N > k_F$) because the former can command a higher output price than the latter, i.e. $p_N > p_F$.[19] In fact, some ethical consumers are willing to pay higher prices for goods and services produced by mission-oriented or non-profit firms.[20] Obviously, the technological or price advantage of the non-profit firm must be sufficiently high as to fully offset the revenue constraint $\alpha^N < 1$.

Workers

Consider a population of agents with unit mass, who differ in two characteristics, ability and intrinsic motivation, that are independently distributed.

Ability takes two values, high and low. A worker characterized by high ability incurs in a low cost of providing a given effort level. Ability is denoted by $\theta \in \{\underline{\theta}, \overline{\theta}\}$ where $\overline{\theta} > \underline{\theta} \geq 1$. A fraction ν of employees has high ability (i.e. a low cost of effort) $\underline{\theta}$, the fraction $1-\nu$ is instead characterized by low ability (i.e. a high cost of effort) $\overline{\theta}$. We will denote by $\Delta\theta$ the difference in ability, whereby $\Delta\theta = \overline{\theta} - \underline{\theta}$.

Intrinsic motivation is continuous and uniformly distributed in the interval $[\underline{\gamma}, \overline{\gamma}]$, with $\overline{\gamma} > \underline{\gamma} \geq 0$. For simplicity, we set $\overline{\gamma} = 1$ and $\underline{\gamma} = 0$, so that $\overline{\gamma} - \underline{\gamma} = 1$, i.e. the support of the distribution is the unit interval. According to Heyes (2005) and Delfgaauw and Dur (2010), we interpret intrinsic motivation as a non-monetary benefit that a worker enjoys when employed by a particular organization, which is unrelated to output produced or effort exerted. In our framework, the premium from intrinsic motivation can only be enjoyed when workers are employed by the non-profit firm, because workers share its mission and observe its commitment in terms of revenue constraint. For instance, health professionals derive, to a certain extent, utility from exerting effort at the non-profit hospital, because only then can they provide treatment to poor and uninsured patients.

When a worker is not hired by any principal, we assume that her utility is zero. If a worker is hired by one principal, her *reservation utility* or *outside option* is endogenous and it depends on the contract offered by the rival principal.

When a worker is hired by the for-profit principal, her utility is given by the salary gained less the cost of effort provision, which depends on the agent's ability type θ. Thus,

$$u_F = w_F - \frac{1}{2}\theta x_F^2$$

In fact, motivated workers do not enjoy any benefit from motivation when hired by the for-profit firm. As a consequence, from the point of view of the for-profit firm, ability is the only relevant workers' characteristic

[19] Glaeser and Shleifer (2001) show that the non-profit status serves as a commitment device to provide softer incentives, which translates into an improvement of the quality of the product sold and into the consumers' higher willingness to pay for the non-profit goods.

[20] Caring consumers are ready to pay higher prices for commodities characterized by some public good attribute (Besley and Ghatak 2007). Think also about parents willing to pay higher tuition fees for religious private schools or ethical investors ready to accept lower interest rates when financing social projects.

because workers with the same ability are the same, irrespective of their level of motivation.[21] Likewise, when a worker is hired by the non-profit firm, her utility takes the form

$$u_N = w_N - \frac{1}{2}\theta x_N^2 + \gamma,$$

where only ability θ is related to effort exertion while motivation γ is not action-oriented. Hence, the premium for intrinsic motivation γ does not directly affect the non-profit firm's output.

Observe that the marginal rate of substitution between effort and wage is given by

$$MRS_{x,w}^i = -\frac{\partial u_i/\partial x_i}{\partial u_i/\partial w_i} = \theta x_i,$$

for each firm $i = F, N$, which is always positive. Indeed, a worker of type θ has preferences over effort-salary pairs which are independent of γ (conditional on being hired by one firm). So all workers' indifference curves have positive slope in the (x, w) plane and the single-crossing property holds for both firms.

Firms' strategic interaction

Following Rochet and Stole (2002), we take the workers' decision to accept the job offered by one firm as given, and we suppose that firms offer incentive-compatible transfer schedules that are conditional on the effort target, i.e. we study non-linear wage schedules $w_i(x_i)$ offered by each firm $i = F, N$. Because a worker of type θ has preferences over effort-salary pairs, which are independent of γ (conditional on being hired by one firm), then we can study the direct revelation mechanism such that each firm offers two incentive-compatible contracts, one for each ability type θ, consisting in an effort target and a wage rate, $\{x_i(\theta), w_i(\theta)\}_{i=F,N}$, and each agent selects the preferred pair. We can thus treat the firms' contract design problem as independent of the workers' choice about which firm to work for. The latter is considered as an indirect mechanism, because no report on γ is required. Given the contracts offered by the two firms, we find the indirect utilities of a worker who truthfully reports her ability type θ and we use them to tackle the worker's self-selection problem, which is determined by motivation γ. This is why, in what follows, it will be more convenient to reason in terms of workers' utility and to focus on contracts of the form $\{x_i(\theta), U_i(\theta)\}_{i=F,N}$.

Let $U_i(\theta)$ denote the *indirect utility* or information rent of an agent of type θ who is hired by firm $i = F, N$, *absent* the benefit accruing from intrinsic motivation. Then

$$U_i = \max_{x_i} w_i(x_i) - \frac{1}{2}\theta x_i^2.$$

Denoting by $x_i(\theta)$ the solution to this program, one can write

$$U_i(\theta) = w_i(x_i(\theta)) - \frac{1}{2}\theta x_i^2(\theta). \tag{2}$$

[21] However, agents with the same ability and different motivation have different outside options.

Given $U_i(\theta)$, it is possible to determine the share of type θ workers employed by each firm, i.e. the probability that type θ workers prefer to be hired by firm i rather than by the rival firm $-i$. Indeed, a worker of type (θ, γ) gets indirect utility $U_F(\theta)$ if she is hired by the for-profit firm, whereas if the same worker is employed by the non-profit firm, her *total* indirect utility becomes

$$\mathcal{U}_N(\theta) = U_N(\theta) + \gamma.$$

Definition 1 ***Indifferent worker.*** *Given ability θ, the worker who is indifferent between working for the non-profit or the for-profit firm is characterized by motivation*

$$\widehat{\gamma}(\theta) = U_F(\theta) - U_N(\theta). \tag{3}$$

Thus, a type (θ, γ) worker strictly prefers to work for the for-profit firm if her motivation falls short of $\widehat{\gamma}(\theta)$, i.e. if $U_N(\theta) + \gamma < U_F(\theta)$; conversely, she strictly prefers to work for the non-profit firm if her motivation exceeds $\widehat{\gamma}(\theta)$ and $U_N(\theta) + \gamma > U_F(\theta)$. For further reference, note that $\widehat{\gamma}(\theta)$ represents the labor donation, i.e. the amount of salary that workers are willing to give up in order to be hired by the non-profit organization.

Given that γ is uniformly distributed on the $[0, 1]$ interval, the share of workers with ability θ who prefer being employed by the for-profit firm is given by

$$\varphi_F(\theta) \equiv \Pr(\gamma < \widehat{\gamma}(\theta)) = U_F(\theta) - U_N(\theta); \tag{4}$$

conversely, the share of agents preferring to be hired by the non-profit firm is

$$\varphi_N(\theta) \equiv \Pr(\gamma \geq \widehat{\gamma}(\theta)) = 1 - (U_F(\theta) - U_N(\theta)). \tag{5}$$

Obviously, in order for both firms to have a positive labor supply by type θ workers, it must be that

$$0 < U_F(\theta) - U_N(\theta) < 1 \Longleftrightarrow \varphi_i(\theta) \in (0, 1) \text{ for each } i = F, N \text{ and each } \theta \in \left\{\underline{\theta}, \overline{\theta}\right\}. \tag{6}$$

This represents the most interesting situation to analyze because no firm is able to attract all the workers of a given ability level. In what follows, we focus our attention precisely on the case in which condition (6) is satisfied and each firm is able to attract both low- and high-ability workers. As we will show, this requires that the two firms be sufficiently similar, or that their marginal revenues be close enough.

Finally, intrinsic motivation generates labor donations from the workers because the inequality $U_N(\theta) < U_F(\theta)$, for every θ, implies that for the same level of effort non-profit workers accept a lower wage. Thus, labor donations represent a benefit for the non-profit firm, despite the fact that they come with a cost, given the non-profit revenue constraints.

Before being able to set up the firms' maximization problem, let us go back to (2) and solve for the wage rate as

$$w_i(\theta) = U_i(\theta) + \frac{1}{2}\theta x_i^2(\theta). \tag{7}$$

One can use expression (7) to eliminate the wage rate from the firm's profits (1). Then, profit margins relative to each type θ worker can be written as

$$\pi_i(\theta) = S_i(\theta) - U_i(\theta) = k_i x_i(\theta) - \frac{1}{2}\theta x_i^2(\theta) - U_i(\theta) \tag{8}$$

where

$$S_i(\theta) \equiv k_i x_i(\theta) - \frac{1}{2}\theta x_i^2(\theta) \tag{9}$$

is the total surplus realized by a worker of type θ providing effort $x_i(\theta)$ for firm i (again, absent the benefit accruing from intrinsic motivation, when $i = N$). Indeed, $S_i(\theta) = \pi_i(\theta) + U_i(\theta)$.

The program of each firm $i = F, N$ is

$$\begin{array}{ll} \max_{x_i(\theta),U_i(\theta)} & E(\pi_i) = \left[\nu\left(k_i x_i(\underline{\theta}) - \frac{1}{2}\underline{\theta}x_i^2(\underline{\theta}) - U_i(\underline{\theta})\right)\varphi_i(\underline{\theta}) + \right. \\ & \left. (1-\nu)\left(k_i x_i(\overline{\theta}) - \frac{1}{2}\overline{\theta}x_i^2(\overline{\theta}) - U_i(\overline{\theta})\right)\varphi_i(\overline{\theta})\right] \end{array} \tag{Pi}$$

Notice that motivation γ does not appear in the program, because it is replaced by the fraction $\varphi_i(\theta)$ of type θ workers being hired by firm $i = F, N$, which in turn depends on the difference between indirect utilities $U_F(\theta) - U_N(\theta)$ (see equations 4 and 5). Moreover, in firm i's program, the (reservation) utility offered by the other firm $U_{-i}(\theta)$ is treated as given but it is endogenous (and dependent on ability only). Thus, firms compete against each other in the utility space: an increase in the utility offered to a given type of worker reduces the firm's payoff when hiring this worker but increases the probability of hiring her.[22] Finally, because ability is not observable by the principals, one has to add to each firm's maximization problem the workers' incentive compatibility constraints. Provided that both firms are able to hire workers with both ability levels, there are two incentive compatibility constraints for each firm: the *downward incentive constraint* (henceforth DIC) meaning that high-ability types should not be attracted by the contract offered to low-ability types and the *upward incentive constraint* (henceforth UIC) meaning that low-ability types are not willing to mimic high-ability workers. For each firm $i = F, N$ such constraints are given by

$$w_i(x_i(\underline{\theta})) - \frac{1}{2}\underline{\theta}x_i^2(\underline{\theta}) \geq w_i\left(x_i(\overline{\theta})\right) - \frac{1}{2}\underline{\theta}x_i^2(\overline{\theta})$$

and

$$w_i\left(x_i(\overline{\theta})\right) - \frac{1}{2}\overline{\theta}x_i^2(\overline{\theta}) \geq w_i(x_i(\underline{\theta})) - \frac{1}{2}\overline{\theta}x_i^2(\underline{\theta}),$$

respectively. Again, observe that these constraints do not depend on γ because motivation enters both sides of each inequality and therefore it cancels out. One can use (7) in order to eliminate wages from the above constraints and rewrite them as a function of effort and utility, so that

$$U_i(\underline{\theta}) \geq U_i(\overline{\theta}) + \frac{1}{2}\left(\overline{\theta} - \underline{\theta}\right)x_i^2(\overline{\theta}) \tag{DIC_i}$$

[22] Notice that the workers' participation constraint does not appear in the firm's program because we study direct mechanisms conditional on the agents' choice to work for each firm.

and

$$U_i\left(\overline{\theta}\right) \geq U_i\left(\underline{\theta}\right) - \frac{1}{2}\left(\overline{\theta} - \underline{\theta}\right) x_i^2\left(\underline{\theta}\right). \tag{UIC_i}$$

In the case of a monopsonistic firm willing to hire workers of unknown ability and with type-independent outside options (the single-principal problem), the relevant constraint is DIC, showing that high-ability workers receive an information rent for being able to mimic low-ability applicants. Given that we are now analyzing a setting with competing principals and type-dependent, endogenous outside options, the UIC constraint can also be relevant. As a consequence, low-ability workers too can receive information rents. Finally, putting DIC_i and UIC_i together yields

$$\frac{1}{2}\left(\overline{\theta} - \underline{\theta}\right) x_i^2\left(\overline{\theta}\right) \leq U_i\left(\underline{\theta}\right) - U_i\left(\overline{\theta}\right) \leq \frac{1}{2}\left(\overline{\theta} - \underline{\theta}\right) x_i^2\left(\underline{\theta}\right)$$

which makes it clear that incentive compatible contracts must satisfy: (i) the *monotonicity* or *implementability condition*

$$x_i\left(\underline{\theta}\right) \geq x_i\left(\overline{\theta}\right), \tag{10}$$

requiring that high-ability workers exert more effort than low-ability types at each firm $i = F, N$; and (ii) condition $U_i\left(\underline{\theta}\right) - U_i\left(\overline{\theta}\right) \geq 0$, requiring that the information rent of high-ability workers be higher than that of low-ability types, for each employer $i = F, N$.

To sum up, each firm $i = F, N$ maximizes its expected profits with respect to the effort level $x_i(\theta)$ and the indirect utility $U_i(\theta)$ set for each type θ worker, taking as given the indirect utility U_{-i} that the rival firm leaves to the workers, subject to the two incentive compatibility constraints DIC and UIC illustrated above. Once the workers' effort levels and utilities are obtained, the related wages w_i are derived using equation (7).

Workers' self-selection

Given the indirect utilities $U_i(\theta)$ set by firms, prospective employees decide which firm to work for according to their level of motivation. This characterizes the workers' self-selection, which is relevant not only under asymmetric information about ability θ, but also when the skills of potential applicants are perfectly observable. Three different sorting patterns of workers to firms are possible.

Definition 2 ***Workers' self-selection.*** *The sorting of workers between the for-profit and the non-profit firm is such that:*

(i) *there is* ***ability neutrality*** *when*

$$\widehat{\gamma}\left(\underline{\theta}\right) = \widehat{\gamma}\left(\overline{\theta}\right) \Longleftrightarrow U_F\left(\underline{\theta}\right) - U_N\left(\underline{\theta}\right) = U_F\left(\overline{\theta}\right) - U_N\left(\overline{\theta}\right); \tag{11}$$

(ii) *there is a* ***negative selection of ability into the non-profit firm*** *when*

$$\widehat{\gamma}\left(\underline{\theta}\right) > \widehat{\gamma}\left(\overline{\theta}\right) \Longleftrightarrow U_F\left(\underline{\theta}\right) - U_N\left(\underline{\theta}\right) > U_F\left(\overline{\theta}\right) - U_N\left(\overline{\theta}\right); \tag{12}$$

(iii) *there is a* ***positive selection of ability into the non-profit firm*** *when*

$$\widehat{\gamma}\left(\underline{\theta}\right) < \widehat{\gamma}\left(\overline{\theta}\right) \Longleftrightarrow U_F\left(\underline{\theta}\right) - U_N\left(\underline{\theta}\right) < U_F\left(\overline{\theta}\right) - U_N\left(\overline{\theta}\right). \tag{13}$$

Ability neutrality captures the situation in which $\varphi_i(\theta)$, i.e. the fraction of workers who self-select into firm $i = F, N$, is constant and does not depend on workers' ability.[23] Negative (respectively, positive) selection into the non-profit firm, instead, means that the fraction of workers attracted by firm N is bigger (resp. smaller) the lower workers' ability, whereas the fraction of workers attracted by firm F is bigger (resp. smaller) the higher workers' ability.[24]

Finally, the timing of the game is as follows. The two firms simultaneously design a menu of contracts of the form $\{x(\theta)_i, U(\theta)_i\}_{i=F,N}$. Workers observe the corresponding non-linear transfer schedule $w_i(x_i)$ for $i = F, N$, select the preferred one and thus choose which firm to work for. Then workers exert their effort level, output is produced, and the contracted wages are paid.

An equilibrium is such that each firm chooses a menu of contracts that maximizes its expected profit, given the contracts offered by the rival principal and given the equilibrium choice of workers. Workers choose the contracts that maximize their utility.

3 The benchmark contracts: full information about ability

Let us first consider the benchmark case in which workers' ability is fully observable, while motivation is the workers' private information. For each type $\theta \in \left\{\underline{\theta}, \overline{\theta}\right\}$, firm $i = F, N$ solves

$$\max_{x_i(\theta), U_i(\theta)} \left(k_i x_i(\theta) - \frac{1}{2}\theta x_i^2(\theta) - U_i(\theta)\right)\varphi_i(\theta) \tag{PBi}$$

taking $U_{-i}(\theta)$, which enters the expression for $\varphi_i(\theta)$, as given. The first-order condition with respect to effort level $x_i(\theta)$ yields

$$x_i^B(\theta) = \frac{k_i}{\theta} = x_i^{FB}(\theta), \tag{14}$$

where $i = F, N$ and where the superindeces B and FB stand for *benchmark* and *first-best*, respectively. In addition, using (14), the first-order conditions with respect to utilities $U_i(\theta)$, which are not symmetric given how function $\varphi_i(\theta)$ is defined, solve for

$$U_F(\theta) = \frac{1}{2}\left(\frac{k_F^2}{2\theta} + U_N(\theta)\right) \quad \text{and} \quad U_N(\theta) = \frac{1}{2}\left(\frac{k_N^2}{2\theta} - (1 - U_F(\theta))\right). \tag{15}$$

[23] When intrinsic motivation is effort-related, both Delfgaauw and Dur (2010, Section 5), under perfect information, and Barigozzi and Burani (2016b), under bidimensional asymmetric information, show that, in equilibrium, sorting is ability neutral.

[24] Delfgaauw and Dur (2010, Section 4) analyze this case when motivation is not effort-related (as in our setting) and firms fully observe workers' characteristics.

These are the reaction functions of the two firms, which characterize the optimal utility left by firm $i = F, N$ to an agent of type θ given the utility $U_{-i}(\theta)$ that this agent receives from the competing firm $-i$. Reaction functions have positive slopes so that utilities can be interpreted as *strategic complements* in this game. In a Nash equilibrium, the levels of utility given by both principals to type θ solve (15) simultaneously so that

$$U_N^B(\theta) = \tfrac{1}{3}\left(\tfrac{k_N^2}{\theta} + \tfrac{k_F^2}{2\theta} - 2\right) \quad \text{and} \quad U_F^B(\theta) = \tfrac{1}{3}\left(\tfrac{k_F^2}{\theta} + \tfrac{k_N^2}{2\theta} - 1\right). \tag{16}$$

Furthermore, we can use expression (3) to obtain the equilibrium value for the marginal worker of type θ, who is indifferent between firms, that is

$$\widehat{\gamma}^B(\theta) = U_F^B(\theta) - U_N^B(\theta) = \frac{1}{3}\left(1 + \frac{k_F^2 - k_N^2}{2\theta}\right),$$

so that the fraction of type θ workers who are hired by firm F is precisely $\varphi_F^B(\theta) = \widehat{\gamma}^B(\theta)$ whereas the fraction of type θ workers that are hired by firm N is $\varphi_N^B(\theta) = 2/3 - \left(k_F^2 - k_N^2\right)/6\theta$. Using (7) one can compute the equilibrium salaries which are such that

$$w_N^B(\theta) = \tfrac{1}{3}\left(\tfrac{5k_N^2 + k_F^2}{2\theta} - 2\right) \quad \text{and} \quad w_F^B(\theta) = \tfrac{1}{3}\left(\tfrac{k_N^2 + 5k_F^2}{2\theta} - 1\right). \tag{17}$$

The Proposition that follows summarizes the results obtained so far.

Proposition 1 ***Benchmark contracts.*** *When ability is observable (and motivation is the workers' private information), the benchmark contracts are a Nash equilibrium of the game in which the non-profit and for-profit firms compete in utility space. The benchmark contracts are such that each firm $i = N, F$ chooses the efficient allocation $x_i^B(\theta) = x_i^{FB}(\theta)$ and leaves to workers utilities given by (16).*

Notice that, at equilibrium, motivated workers employed by the non-profit organization, not only enjoy utility $U_N^B(\theta)$ but also their motivational premium, so that their total indirect utility becomes $\mathcal{U}_N(\theta) = U_N^B(\theta) + \gamma$.[25]

At equilibrium, how do workers characterized by different levels of ability sort between the two firms? It depends on the difference between the firms' marginal revenues, i.e. k_F and k_N.

(i) Let us first consider the case in which both firms have the same marginal revenues so that $k_F = k_N = k$. Then

$$U_N^B(\theta) = \frac{\left(3k^2 - 4\theta\right)}{6\theta} \quad \text{and} \quad U_F^B(\theta) = \frac{\left(3k^2 - 2\theta\right)}{6\theta}$$

with $U_F^B(\theta) > U_N^B(\theta)$. Moreover,

$$\widehat{\gamma}^B(\theta) = U_F^B(\theta) - U_N^B(\theta) = \frac{1}{3},$$

[25] Also observe that, at the equilibrium under full information with respect to both ability and motivation, workers employed by the non-profit firm receive a total indirect utility equal to $U_N^B(\theta)$ because the firm is able to appropriate their motivational premium by paying out only $w_N^B(\theta) - \gamma$.

whereby the indifferent worker has motivation $\gamma = \frac{1}{3}$ independently of her ability. All workers with motivation higher than $\frac{1}{3}$ prefer to work for the non-profit while all workers with motivation lower than $\frac{1}{3}$ prefer to apply at the for-profit firm. The premium γ earned by motivated workers ensures the non-profit firm a labor supply that is twice the one of the for-profit firm.[26]

(ii) Consider now the case in which $k_F > k_N$ so that the for-profit firm has a competitive advantage with respect to the non-profit firm. The indifferent worker with high ability has higher motivation than the low ability one, i.e. $\widehat{\gamma}^B(\underline{\theta}) > \widehat{\gamma}^B(\overline{\theta})$. This means that $\widehat{\gamma}(\theta)$ is decreasing in θ, so that the share of low-ability workers hired by the non-profit firm is larger than the share of high-ability workers. Then, a negative selection of ability into the non-profit firm realizes. [27]

(iii) To conclude, consider the case in which the non-profit firm has a competitive advantage, despite its revenue constraints, $k_F < k_N$. Now $\widehat{\gamma}^B(\underline{\theta}) < \widehat{\gamma}^B(\overline{\theta})$ holds, meaning that $\widehat{\gamma}(\theta)$ is increasing in θ, and we observe a positive selection of ability into the non-profit firm.[28]

The Proposition below focuses on the consequences that the difference in revenues between firms has on workers' self-selection.

Proposition 2 ***Workers' sorting patterns at the benchmark contracts.*** *When ability is observable (and motivation is the workers' private information), benchmark contracts are such that the sorting of workers between firms only depends on the difference in firm's marginal revenues, i.e. on which firm has a competitive advantage over the other:* (i) *if* $k_F = k_N$ *there is ability-neutrality and* $\widehat{\gamma}^B(\underline{\theta}) = \widehat{\gamma}^B(\overline{\theta})$; (ii) *if* $k_F > k_N$ *there is a negative selection of ability into firm* N *and* $\widehat{\gamma}^B(\underline{\theta}) > \widehat{\gamma}^B(\overline{\theta})$ *holds; and* (iii) *if* $k_F < k_N$ *there is a positive selection of ability into firm* N *and* $\widehat{\gamma}^B(\underline{\theta}) < \widehat{\gamma}^B(\overline{\theta})$ *holds.*

[26] Using condition (6) one can check that all workers' types supply a positive amount of labour to both firms if and only if marginal revenues are sufficiently high that $U_N^B(\theta_2) > 0$, or else if and only if $k > \sqrt{\frac{4\theta_2}{3}}$.

[27] Again, from condition (6), an interior solution exists if a positive mass of each type of workers is applying to each firm, that is if $\varphi_F(\theta_1) = \widehat{\gamma}^B(\theta_1) < 1$, i.e. if $k_F^2 - k_N^2 < 4\theta_1$, or else if the difference in firms' revenues is not too high. When the previous condition is not satisfied, then the non-profit firm only hires low-ability workers, given that all high-ability applicants self-select into the for-profit firm.

[28] The non-profit firm always hires a positive mass of both high- and low-ability workers, whereas it might be the case that the for-profit firm only hires low-ability workers. A positive mass of low-ability workers for the for-profit firm, now requires that $k_N^2 - k_F^2 < 2\theta_1$: again, the difference in revenues must not be too high.

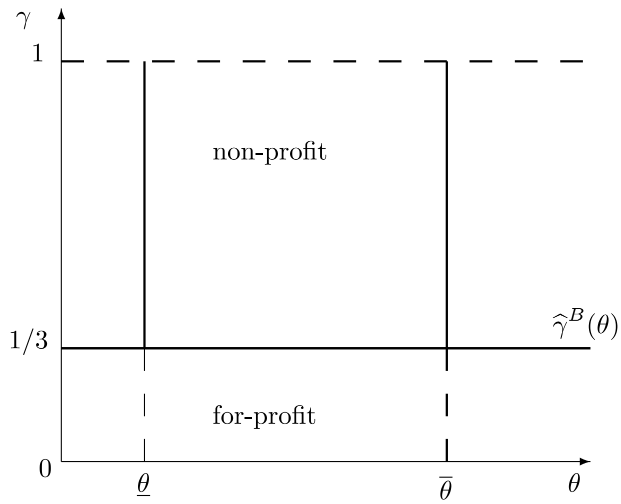


Figure 1(a). Ability-neutrality: $k_F = k_N$

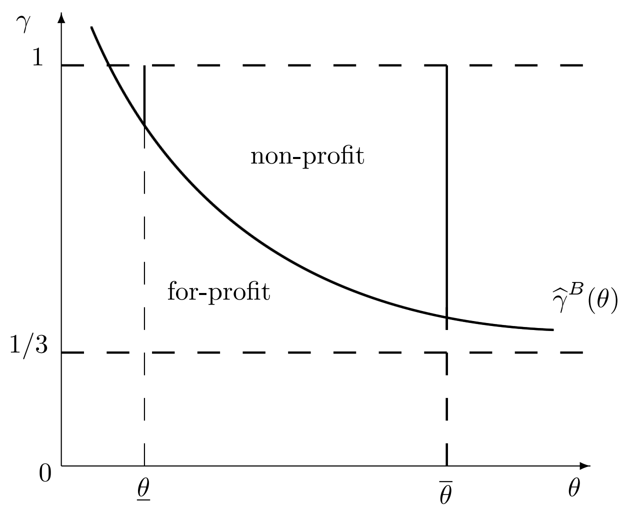


Figure 1(b). Adverse selection of ability into the non-profit firm: $k_F > k_N$

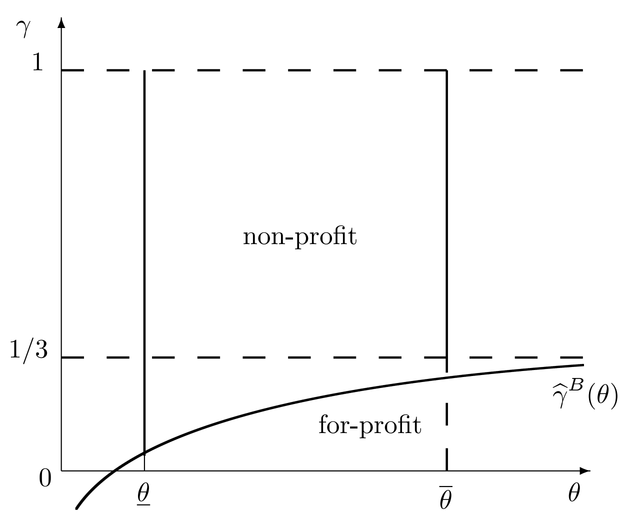


Figure 1(c). Propitious selection of ability into the non-profit firm: $k_N > k_F$

Before moving to the case in which ability is private information, we would like to emphasize the following. Independently of the sign of the difference in marginal revenues between firms, an interior solution exists, meaning that both firms are able to hire workers of each skill level, provided that the difference in revenues be sufficiently small. Otherwise, the advantaged firm is able to hire all workers of a given skill level. In any case, full market segmentation according to skills never occurs, i.e. it is never the case that all workers of a given skill level prefer to work for one firm, whereas all workers with the other skill level prefer to be hired by the rival firm.

4 Screening for ability and incentive contracts

Let us now solve the complete problem in which neither ability nor motivation are observable. This requires taking into account DIC and UIC constraints. Let us first state some preliminary results. The first set of results provides the conditions under which the benchmark contracts analysed in Section 3 represent full-fledged optimal contracts under asymmetric information about both workers' ability and motivation. When those conditions are not fulfilled, we then provide further results that help reduce the set of relevant incentive constraints to be considered for each firm.

Following the arguments developed by Biglaiser and Mezzetti (2000), we first provide a sufficient condition under which the benchmark contracts are incentive compatible for both firms. Then benchmark contracts are the solution to the firms' problems also under asymmetric information about workers' ability. Under this condition, for each firm $i = F, N$, if agents of each type $\theta \in \{\underline{\theta}, \overline{\theta}\}$ exert the first-best effort level $x_i^{FB}(\theta)$ and are compensated according to $w_i^B(\theta)$ specified by (17), then low-ability workers do not want to mimic high-ability agents and strictly prefers the contract $\left(x_i^{FB}(\overline{\theta}), w_i^B(\overline{\theta})\right)$ to the contract $\left(x_i^{FB}(\underline{\theta}), w_i^B(\underline{\theta})\right)$. But the reverse is also true: high-ability workers do not want to mimic low-ability agents. Hence, both incentive constraints are slack for both firms at the benchmark contracts and each firm's problem can be treated as two independent problems, one for each ability level, since the presence of types $\underline{\theta}$ does not influence the optimal contract that firm i offers to types $\overline{\theta}$ and vice-versa.

Lemma 1 *At the benchmark contracts, all incentive constraints are slack for both the for-profit and the non-profit firm if*

$$\frac{(k_F + k_N)\,|k_F - k_N|}{3 \min\{k_F^2; k_N^2\}} < \frac{\overline{\theta} - \underline{\theta}}{\underline{\theta}} \tag{18}$$

holds.

Proof. See Appendix A.1. ■

The Proposition that follows is an immediate consequence of Lemma 1.

Proposition 3 ***Incentive compatible benchmark contracts.*** *Suppose that neither ability nor motivation is observable. When condition (18) is satisfied, then optimal incentive contracts coincide with the* *benchmark contracts: for all* $\theta \in \left\{\underline{\theta}, \overline{\theta}\right\}$ *, both firms* $i = F, N$ *ask workers to exert first-best effort levels* $x_i^*(\theta) = x_i^B(\theta) = x_i^{FB}(\theta)$ *and provide compensation schemes* $w_i^*(\theta) = w_i^B(\theta)$*.*

Lemma 1 provides the condition under which competition between two non-identical firms leads to an efficient allocation.[29] Notice that this efficiency result is more likely to be attained when the difference in firms' types, namely the difference in firms' marginal revenues $|k_F - k_N|$, is sufficiently low relative to the difference in workers' types, i.e. the difference in the costs of effort provision $\overline{\theta} - \underline{\theta}$. When this is not the case, i.e. when condition (18) is not satisfied, it means that the benchmark contracts of at least one firm (the disadvantaged one) are no longer incentive compatible. Then, the following might happen.

Lemma 2 (i) *If condition (18) fails to hold but condition*

$$\frac{(k_F + k_N)|k_F - k_N|}{\min\{k_F^2; k_N^2\} + 2\max\{k_F^2; k_N^2\}} < \frac{\overline{\theta} - \underline{\theta}}{\underline{\theta}} \leq \frac{(k_F + k_N)|k_F - k_N|}{3\min\{k_F^2; k_N^2\}} \tag{19}$$

is satisfied, then the benchmark contracts offered by the firm with the competitive advantage are still incentive compatible, whereas UIC *might fail to be satisfied by the benchmark contracts of the disadvantaged firm.* (ii) *If condition (19) fails to hold but condition*

$$\frac{\overline{\theta} - \underline{\theta}}{\underline{\theta}} \leq \frac{(k_F + k_N)|k_F - k_N|}{\min\{k_F^2; k_N^2\} + 2\max\{k_F^2; k_N^2\}} \tag{20}$$

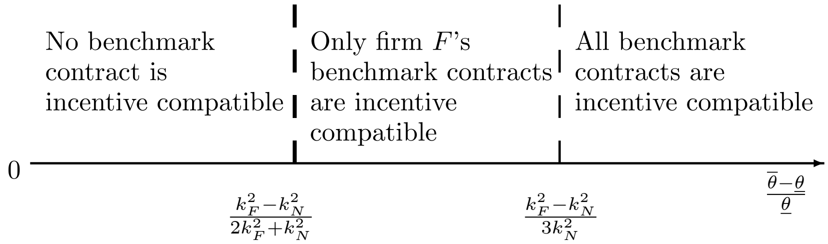
is satisfied, then neither firm's benchmark contracts are incentive compatible and UIC *might fail to be satisfied by the benchmark contracts of the disadvantaged firm and* DIC *might fail to be satisfied by the benchmark contracts of the advantaged firm.*

Proof. See Appendix A.1. ■

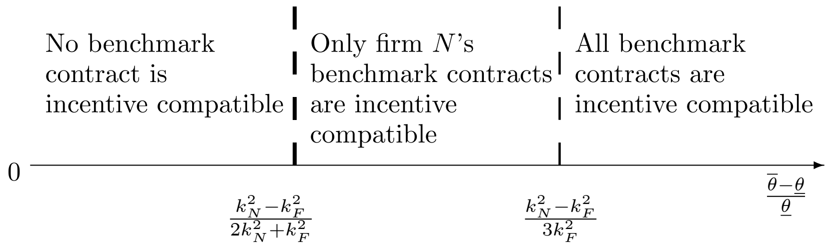
In words, Lemma 2 states that, when the sufficient condition (18) is not met, UIC is binding first for the disadvantaged firm and then DIC is binding for the advantaged firm.

Figure 2 represents the relevant thresholds appearing in Lemmata 1 and 2, specifying which firm has a competitive advantage relative to the rival.

[29] Notice that in Rochet and Stole (2002), due to the symmetry between firms, no incentive constraint can ever be binding. Therefore in Rochet and Stole (2002) and in and Armstrong and Vickers (2001) as well, optimal contracts always consist in efficient allocations and cost-plus-fixed-fee pricing.



(a) $k_F > k_N$



(b) $k_N > k_F$

Figure 2. Incentive compatible benchmark contracts

Finally, provided that the benchmark contracts no longer represent the optimal incentive contracts, Lemma 3 specifies which are the incentive compatibility constraints that each firm can neglect, according to the sorting pattern of workers into firms.[30]

Lemma 3 *(i) When there is ability-neutrality and $\widehat{\gamma}(\underline{\theta}) = \widehat{\gamma}(\overline{\theta})$ holds, then neither DIC nor UIC can be binding for either firm. (ii) When there is a negative selection of ability into firm N and $\widehat{\gamma}(\underline{\theta}) > \widehat{\gamma}(\overline{\theta})$ holds, then neither UIC_F nor DIC_N can be binding. (iii) When there is a positive selection of ability into firm N and $\widehat{\gamma}(\underline{\theta}) < \widehat{\gamma}(\overline{\theta})$ holds, then neither DIC_F nor UIC_N can be binding.*

Proof. See Appendix A.2. ■

The cases (*ii*) and (*iii*) presented in Lemma 3 are outlined in more detail in the subsections that follow.

Before beginning the analysis, let us introduce a restriction on skill levels which is needed in order to ensure that firms make nonnegative profit margins on all ability types (see Appendix A.4 for more details).

[30] These results stand in contrast with Biglaiser and Mezzetti (2000), where it is hown that UIC can never be binding.

Assumption 1 *The difference in ability is sufficiently low so that* $2\underline{\theta} > \overline{\theta} > \underline{\theta} \geq 1$ *holds.*

Thus, provided that Assumption 1 holds, each firm is able to hire a positive mass of workers of each ability-type (and this allows us to focus on the intensive rather than the extensive margin).

4.1 Negative selection of ability into the non-profit firm

4.1.1 UIC binds for the non-profit firm

Take the case in which $k_N < k_F$ and condition (18) fails to hold but condition (19) is satisfied. This is the case in which UIC_N might bind while all incentives constraints are slack for firm F. Now, the program for firm F is the unconstrained (PF) whereas the problem for firm N is (PN) subject to UIC_N binding that is

$$U_N\left(\overline{\theta}\right) = U_N\left(\underline{\theta}\right) - \frac{1}{2}\left(\overline{\theta} - \underline{\theta}\right) x_N^2\left(\underline{\theta}\right).$$

We refer the reader to Appendix A.3 for a detailed analysis of this case.

The proposition that follows provides the most important qualitative results, focusing on allocative distortions and on informational rents (i.e. on how utilities left to the different types of workers change with respect to the benchmark contracts). The information about optimal wages is provided later on, in Section 5.

Proposition 4 ***Optimal incentive contracts when*** UIC_N ***binds.*** *When* $k_N < k_F$ *and condition (18) is not satisfied whereas condition (19) holds, optimal contracts are such that:* (i) *the for-profit firm sets effort levels at the first-best, i.e.* $x_F^*(\theta) = x_F^{FB}(\theta)$ *for each* $\theta \in \left\{\underline{\theta}, \overline{\theta}\right\}$, *and the non-profit firm sets an efficient allocation for low-ability workers, i.e.* $x_N\left(\overline{\theta}\right) = x_N^{FB}\left(\overline{\theta}\right)$, *whereas it distorts high-ability workers' effort upwards, i.e.* $x_N^*\left(\underline{\theta}\right) > x_N^{FB}\left(\underline{\theta}\right)$, *with* $x_N^*(\theta) < x_F^*(\theta)$ *for each* $\theta \in \left\{\underline{\theta}, \overline{\theta}\right\}$; *and* (ii) *the utilities of high-ability workers are lower whereas the utilities of low-ability workers are higher than at the benchmark, i.e.* $U_i^*\left(\underline{\theta}\right) < U_i^B\left(\underline{\theta}\right)$ *while* $U_i^*\left(\overline{\theta}\right) > U_i^B\left(\overline{\theta}\right)$ *for each* $i = F, N$.

Proof. See Appendix A.3. ■

With respect to the benchmark contracts, what changes is that the difference in ability between types decreases. This facilitates mimicking between agents with different ability levels. In particular, the contract offered by firm N to high-ability workers, i.e. types $\underline{\theta}$, becomes attractive for low-ability applicant, i.e. types $\overline{\theta}$.[31] Thus, firm N is forced to distort effort of high-ability types $\underline{\theta}$ upwards in order

[31] In order to give some more intuition, consider the following: firm F has a competitive advantage over firm N, therefore the former is able to leave to its applicants a high utility, namely a high outside option. Because the difference in ability levels is sufficiently high in this case, it means that the outside option left by firm F to high-ability workers is not only high in absolute terms, but also relative to the outside option left by the same firm F to low-ability types. Then , firm N has to meet these high offers of the competitor, and it is bound to leave to high-ability workers a high utility. This is why micking from low-ability types becomes attractive.

to make mimicking less attractive and, at the same time, to give information rents to low-ability types $\overline{\theta}$ whose utility increases, whereby $U_N^*\left(\overline{\theta}\right) > U_N^B\left(\overline{\theta}\right)$ (see inequality 41 in Appendix A.3). Since utilities are strategic complements, an increase in $U_N\left(\overline{\theta}\right)$ also leads to an increase in $U_F\left(\overline{\theta}\right)$, although the rate of change of $U_F\left(\overline{\theta}\right)$ is half the rate of change of $U_N\left(\overline{\theta}\right)$. Then the probability of type $\overline{\theta}$ workers self-selecting into firm N increases as well with respect to the benchmark contract and the effect of negative selection of ability is reinforced. In fact, the difference $U_F\left(\overline{\theta}\right) - U_N\left(\overline{\theta}\right)$ shrinks with respect to the benchmark and $\widehat{\gamma}^*\left(\overline{\theta}\right) < \widehat{\gamma}^B\left(\overline{\theta}\right)$.

Moreover, the first-order conditions also imply that $U_N^*\left(\underline{\theta}\right) < U_N^B\left(\underline{\theta}\right)$. Again, the strategic complementarity in utilities leads to conclude that also $U_F^*\left(\underline{\theta}\right)$ decreases with respect to the benchmark case, but less than $U_N^*\left(\underline{\theta}\right)$. Therefore the difference $U_F\left(\underline{\theta}\right) - U_N\left(\underline{\theta}\right)$ increases with respect to the benchmark and we have $\widehat{\gamma}^*\left(\underline{\theta}\right) > \widehat{\gamma}^B\left(\underline{\theta}\right)$.

4.1.2 *UIC* binds for the non-profit firm and *DIC* binds for the for-profit firm

Consider the case in which $k_F > k_N$ and both conditions (18) and (19) fail to hold so that neither firm can treat its contract offered to low-ability agents as independent of the contract offered to high-ability agents and vice-versa. In particular, UIC binds for firm N while DIC binds for firm F. Now, the program of firm N is (PN) subject to UIC_N binding

$$U_N\left(\overline{\theta}\right) = U_N\left(\underline{\theta}\right) - \frac{1}{2}\left(\overline{\theta} - \underline{\theta}\right) x_N^2\left(\underline{\theta}\right),$$

as in the preceding case, whereas the program of firm F is (PF) subject to DIC_F binding

$$U_F\left(\underline{\theta}\right) = U_F\left(\overline{\theta}\right) + \frac{1}{2}\left(\overline{\theta} - \underline{\theta}\right) x_F^2\left(\overline{\theta}\right).$$

Again, the Proposition that follows highlights the most relevant qualitative features of this equilibrium. We refer the reader to Appendix A.5 for the detailed analysis of the system of first-order conditions that characterize the solution in this case.

Proposition 5 ***Optimal incentive contracts when*** UIC_N ***and*** DIC_F ***bind.*** *When* $k_N < k_F$ *and neither condition (18) nor (19) is satisfied while condition (20) holds, optimal contract are such that: (i) the for-profit firm sets an efficient allocation for high-ability workers, i.e.* $x_F^*\left(\underline{\theta}\right) = x_F^{FB}\left(\underline{\theta}\right)$*, whereas it distorts downward the effort of low-ability workers, i.e.* $x_F^*\left(\overline{\theta}\right) < x_F^{FB}\left(\overline{\theta}\right)$*; the non-profit firm sets an efficient allocation for low-ability workers, i.e.* $x_N^*\left(\overline{\theta}\right) = x_N^{FB}\left(\overline{\theta}\right)$*, whereas it distorts upwards the effort of high-ability workers, i.e.* $x_N^*\left(\underline{\theta}\right) > x_N^{FB}\left(\underline{\theta}\right)$*; therefore effort levels are such that* $x_N^*\left(\overline{\theta}\right) < x_N^*\left(\underline{\theta}\right) < x_F^*\left(\overline{\theta}\right) < x_F^*\left(\underline{\theta}\right)$*; (ii) utilities* $U_N^*\left(\overline{\theta}\right)$ *and* $U_F^*\left(\underline{\theta}\right)$ *are higher whereas utilities* $U_N^*\left(\underline{\theta}\right)$ *and* $U_F^*\left(\underline{\theta}\right)$ *are lower than at the benchmark contracts.*

Proof. See Appendix A.5. ■

What is new to this case is that the contract offered by firm F to low-ability workers becomes attractive for high-ability potential applicants. Firm F prevents high-ability workers from mimicking low-ability types by distorting the effort level required from low-skilled workers downwards and increasing the rents left to high-skilled applicants, so that $U_F^*(\underline{\theta})$ necessarily increases with respect to the benchmark contracts. Moreover, since first-order conditions require that the utilities that the same firm leaves to different types of agents move in opposite directions (see Appendix A.5), an increase in $U_F^*(\underline{\theta})$ is accompanied by a decrease in $U_F^*\left(\overline{\theta}\right)$. Analogously, the increase in $U_N^*\left(\overline{\theta}\right)$ that firm N implements in order to discourage low-skilled workers from mimicking high-skilled ones, goes hand in hand with a decrease in $U_N^*(\underline{\theta})$. Thus, the difference with the previous case, in which only UIC_N is binding, stems from the fact that strategic complementarity is no longer relevant and that the utilities that the two firms offer to the same type of worker no longer vary in the same direction: now they move in opposite directions. This has clear-cut implications for the difference $U_F^*(\theta) - U_N^*(\theta) = \widehat{\gamma}^*(\theta)$. Indeed, $\widehat{\gamma}^*(\underline{\theta}) > \widehat{\gamma}^B(\underline{\theta})$ and $\widehat{\gamma}^*\left(\overline{\theta}\right) < \widehat{\gamma}^B\left(\overline{\theta}\right)$ both hold (as when only UIC_N is binding) because the changes in utilities with respect to the benchmark reinforce each other. This implies that incentive contracts exacerbate the negative selection effect into the non-profit firm which can already be observed at the benchmark contracts.

In the next subsection, we consider the symmetric case of positive selection. The uninterested reader might skip this part and move directly to Section 4.3.

4.2 Positive selection of ability into the non-profit firm

The analysis of these cases is symmetric to the one in subsection 4.1, therefore we refer the interested reader to Appendices A.6 and A.7 and we only state here the main qualitative results.

4.2.1 UIC binds for the for-profit firm

Consider the case in which $k_F < k_N$ and in which condition (18) fails to hold but condition (19) is satisfied. Then UIC_F might bind while all incentives constraints are slack for firm N. In particular, the program for firm N is the unconstrained (PN) whereas the problem for firm F is (PF) subject to UIC_F binding

$$U_F\left(\overline{\theta}\right) = U_F(\underline{\theta}) - \frac{1}{2}\left(\overline{\theta} - \underline{\theta}\right) x_F^2(\underline{\theta}).$$

The optimal contracts are characterized in the Proposition that follows.

Proposition 6 ***Optimal incentive contracts when*** UIC_F ***binds.*** *When $k_N > k_F$ and condition (18) is not satisfied whereas condition (19) holds, optimal contracts are such that:* (i) *the non-profit firm sets effort levels at the first-best, i.e. $x_N^*(\theta) = x_N^{FB}(\theta)$ for each $\theta \in \left\{\underline{\theta}, \overline{\theta}\right\}$, and the for-profit firm sets an efficient allocation for low-ability workers, i.e. $x_F\left(\overline{\theta}\right) = x_F^{FB}\left(\overline{\theta}\right)$, whereas it distorts high-ability workers' effort upwards, i.e. $x_F^*(\underline{\theta}) > x_F^{FB}(\underline{\theta})$, with $x_F^*(\theta) > x_N^*(\theta)$ for each $\theta \in \left\{\underline{\theta}, \overline{\theta}\right\}$;* (ii) *the utilities offered* *to low-ability workers are higher whereas the utilities offered to high-ability workers are lower than at the benchmark, i.e.* $U_i^*\left(\overline{\theta}\right) > U_i^B\left(\overline{\theta}\right)$ *while* $U_i^*\left(\underline{\theta}\right) < U_i^B\left(\underline{\theta}\right)$ *for each* $i = F, N$.

Proof. See Appendix A.6. ∎

4.2.2 UIC binds for the for-profit firm and DIC binds for the non-profit firm

Finally, consider the case in which $k_N > k_F$ and both conditions (18) and (19) fail to hold, so that neither firm can treat its contract offered to low-ability agents as independent of the contract offered to high-ability agents and vice-versa. Now, the program of firm F is (PF) subject to UIC_F binding

$$U_F\left(\overline{\theta}\right) = U_F\left(\underline{\theta}\right) - \frac{1}{2}\left(\overline{\theta} - \underline{\theta}\right) x_F^2\left(\underline{\theta}\right),$$

as in the preceding case, whereas the program of firm N is (PN) subject to DIC_N binding

$$U_N\left(\underline{\theta}\right) = U_N\left(\overline{\theta}\right) + \frac{1}{2}\left(\overline{\theta} - \underline{\theta}\right) x_N^2\left(\overline{\theta}\right).$$

Again, the Proposition that follows highlights the most relevant qualitative features of this equilibrium.

Proposition 7 ***Optimal incentive contracts when*** UIC_F ***and*** DIC_N ***bind.*** *When* $k_N > k_F$ *and neither condition (18) nor condition (19) holds, optimal contracts are such that:* (i) *the non-profit firm sets an efficient allocation for high-ability workers, i.e.* $x_N^*\left(\underline{\theta}\right) = x_N^{FB}\left(\underline{\theta}\right)$, *whereas it distorts downward the effort of low-ability workers, i.e.* $x_N^*\left(\overline{\theta}\right) < x_N^{FB}\left(\overline{\theta}\right)$; *the for-profit firm sets an efficient allocation for low-ability workers, i.e.* $x_F^*\left(\overline{\theta}\right) = x_F^{FB}\left(\overline{\theta}\right)$, *whereas it distorts upward the effort of high-ability workers, i.e.* $x_F^*\left(\underline{\theta}\right) > x_F^{FB}\left(\underline{\theta}\right)$; *therefore effort levels are such that* $x_F^*\left(\overline{\theta}\right) < x_F^*\left(\underline{\theta}\right) < x_N^*\left(\overline{\theta}\right) < x_N^*\left(\underline{\theta}\right)$; (ii) *utilities* $U_F^*\left(\overline{\theta}\right)$ *and* $U_N^*\left(\underline{\theta}\right)$ *are higher whereas utilities* $U_F^*\left(\underline{\theta}\right)$ *and* $U_N^*\left(\overline{\theta}\right)$ *are lower than at the benchmark contracts.*

Proof. See Appendix A.7. ∎

Before moving to the next section and analyse optimal wages, let us consider the sorting pattern of workers to firms.

4.3 Incentive contracts and workers' sorting patterns

When the non-profit firm has a competitive disadvantage with respect to the for-profit firm , i.e. $k_N < k_F$, and incentive contracts are in place, there is negative selection of ability into the non-profit firm and we find that $\widehat{\gamma}^*\left(\underline{\theta}\right) > \widehat{\gamma}^B\left(\underline{\theta}\right)$ and $\widehat{\gamma}^*\left(\overline{\theta}\right) < \widehat{\gamma}^B\left(\overline{\theta}\right)$. In other words, the negative selection effect is exacerbated with respect to the benchmark contracts. This implies that the labor supply from high-ability workers faced by the non-profit firm (respectively the for-profit firm) decreases (resp. increases) relative to the benchmark, whereas the labor supply from low-ability workers faced by the non-profit firm (respectively the for-profit firm) increases (resp. decreases) relative to the benchmark.

This stands in contrast to the case in which $k_N > k_F$ holds and there is positive selection into the non-profit firm. Then both $\widehat{\gamma}^* (\underline{\theta}) < \widehat{\gamma}^B (\underline{\theta})$ and $\widehat{\gamma}^* (\overline{\theta}) > \widehat{\gamma}^B (\overline{\theta})$ are true. This means that when the non-profit firm has a competitive advantage over the for-profit firm and incentive contracts are needed, then the positive selection effect is reinforced with respect to the benchmark contracts. Then, the labor supply from high-ability workers faced by the non-profit firm (respectively the for-profit firm) increases (resp. decreases) relative to the benchmark, whereas the labor supply from low-ability workers faced by the non-profit firm (respectively the for-profit firm) decreases (resp. increases) relative to the benchmark contacts.

It is then possible to bunch both cases of negative and positive selection and derive a general statement about how workers' self-selection into the non-profit and the for-profit firm changes when the benchmark contracts are no longer incentive compatible.

Proposition 8 ***Workers' sorting patterns at the incentive contracts.*** *When condition (18) is not satisfied and incentive contracts are in place, the selection effects of ability are more pronounced (i.e. the function $\widehat{\gamma} (\theta)$ is steeper) than at the benchmark contracts.*

The above Proposition suggests that each firm designs its incentive contracts in such a way as to make the sorting pattern of workers even more favorable to itself. Indeed, when UIC is binding for one firm, its profit margins are higher for low-ability than for high-ability workers.[32] Therefore the disadvantaged firm, whose UIC is binding, is better-off the higher the fraction of low-ability workers that it is able to hire and the lower the fraction of high-ability workers that it captures. The opposite happens when DIC is binding for the firm with a competitive advantage: its profit margins are higher for high-ability than for low-ability workers, therefore this firm is better-off if it succeeds in hiring an increasing fraction of high-ability workers and a decreasing share of low-ability workers. Therefore, it becomes relatively more convenient for the advantaged firm to attract high-ability workers and for the disadvantaged firm to attract low-ability applicants. That is why the selection effects of ability are exacerbated at the incentive contracts.

Finally, Proposition 8 hints at the possibility that incentive contracts might have a sort of exclusionary effect. Indeed, starting from an interior solution at the benchmark contracts, one might observe that under the incentive contracts the following happens: the supply of high-skilled labor might vanish either for firm N under negative selection or for firm F under positive selection. In sum, incentive contracts might drive the supply of high-skilled labor to zero for the disadvantaged firm.

[32] See the preliminary Result 1 contained in the proof of Lemma 3 in Appendix A.2.

5 Wage differentials

In this section, we analyze the wage differential, i.e. $sign\left(w_F\left(\theta\right)-w_N\left(\theta\right)\right)$, that characterizes the contracts offered in equilibrium by the two firms. In particular, we study whether a worker with given ability θ is paid more by the for-profit or by the non-profit firm. Interestingly, our results concern not only the sign of the wage differential but also its composition: we are able to disentangle the effect of labor donations to the non-profit firm stemming from workers' motivation from the effect of the negative or positive selection into the non-profit firm (which depends upon which firm holds a competitive advantage relative to the other).

Fixing θ, let us then consider when $sign\left(w_F\left(\theta\right)-w_N\left(\theta\right)\right)>0$ holds, i.e. when a wage penalty for non-profit workers is in place. Using expression (7) for the wage rate, the wage differential can be rewritten as

$$U_F\left(\theta\right)+\frac{1}{2}\theta x_F^2\left(\theta\right)-\left(U_N\left(\theta\right)+\frac{1}{2}\theta x_N^2\left(\theta\right)\right)>0,$$

or, rearranging, as

$$\underbrace{U_F\left(\theta\right)-U_N\left(\theta\right)}_{\substack{\equiv\widehat{\gamma}(\theta)\\ \text{labor donation (+)}}}+\underbrace{\frac{1}{2}\theta\left(x_F^2\left(\theta\right)-x_N^2\left(\theta\right)\right)}_{\text{negative selection (+) / positive selection (-)}}>0, \tag{21}$$

for each $\theta\in\left\{\underline{\theta},\overline{\theta}\right\}$. Expression (21) above contains two terms. The first one is the difference between worker θ's utility at the two firms, and is always positive because it corresponds to the level of motivation of the indifferent worker, i.e. $\widehat{\gamma}\left(\theta\right)$, which is strictly positive at an interior solutions, such that the share of workers with ability θ applying to any firm $i=F,N$ is always positive. Thus, the first term represents the labor donation and corresponds to the amount of salary that sufficiently motivated workers (that is workers with motivation $\gamma\geq\widehat{\gamma}\left(\theta\right)$) are willing to give up in order to be hired by the non-profit firm. The second term is the difference between the squared levels of effort set by the two firms, where $x_F\left(\theta\right)>x_N\left(\theta\right)$ when the selection into the non-profit firm is negative (because $k_F>k_N$) or $x_F\left(\theta\right)<x_N\left(\theta\right)$ when the selection into the non-profit firm is positive (because $k_N>k_F$). Thus, the second term in expression (21) reflects the impact of the selection of workers into firms, which in turn depends on which firm has a competitive advantage over the other.

Notice that the unique instance in which the wage differential turns out to be negative, so that $w_F\left(\theta\right)<w_N\left(\theta\right)$ and non-profit employees experience a wage premium, is when not only is there positive selection into the non-profit firm but also when the selection effect is strong enough to offset the labor donation effect. This case requires that the competitive advantage of the non-profit firm be relevant enough.

Wage differentials not only arise at the optimal incentive contracts: they are already in place at the benchmark contracts. So, in what follows, let us distinguish between the two cases.

5.1 Benchmark contracts

Let us consider first the benchmark contracts. In particular, suppose that motivation is the worker's private information and that either ability is observable or ability is not observable but the sufficient condition (18) is satisfied.

Proposition 9 ***Wage differentials at the benchmark contracts.*** *At the benchmark contracts, a wage penalty for non-profit employees is in place, i.e.* $w_F^B(\theta) > w_N^B(\theta)$ *holds for all* $\theta \in \{\underline{\theta}, \overline{\theta}\}$*, unless there is positive selection of ability into firm* N *and the competitive advantage of firm* N *is sufficiently high that* $k_N^2 - k_F^2 \geq \frac{\overline{\theta}}{2}$.

More specifically, consider the expressions for wages given by equation (17). Suppose first that no firm has a competitive advantage over the other so that $k_N = k_F = k$. Then, it is easy to see that wages offered by the two firms are always such that $w_F^B(\theta) > w_N^B(\theta)$ for every θ. For each level of workers' ability, a wage differential favoring workers employed at the for-profit firm is in place. Indeed, the for-profit firm asks its employees to provide the same first-best effort that is required by the non-profit firm, but in exchange for a higher salary. Here the wage differential is purely compensating because it originates uniquely from labor donations.

When, instead, firm F holds a competitive advantage and $k_N < k_F$, the selection effect is always positive. A wage premium in favour of for-profit workers exists also in this case, but workers are now asked to exert a higher effort at the for-profit than at the non-profit firm.

Finally, when $k_N > k_F$ and the selection effect is negative, the wage differential may have a different sign according to the magnitude of the difference in firms' revenues. In particular, a wage premium favoring for-profit workers still exists provided that $k_N^2 - k_F^2 < \frac{\underline{\theta}}{2}$. Alternatively, if $\frac{\underline{\theta}}{2} \leq k_N^2 - k_F^2 < \frac{\overline{\theta}}{2}$, then high-ability workers earn more when they are employed by the non-profit than by the for-profit firm, whereas low-ability workers earn more when they are employed by the for-profit than by the non-profit firm. This result is indeed peculiar given that the wage differential changes its sign according to the ability of the workers. Finally, if $k_N^2 - k_F^2 \geq \frac{\overline{\theta}}{2}$, then all workers get a wage premium when hired by the non-profit firm. Thus, when the non-profit firm has a competitive advantage over the for-profit rival, both wage premia and penalties can be observed. If the difference in marginal revenues is not too high, then non-profit workers exert more effort but are paid less than for-profit employees, given their ability. When instead the advantage of the non-profit firm is sufficiently important, then the non-profit workers' higher effort is rewarded with a higher salary.

5.2 Incentive contracts

When neither ability nor motivation is observable and the sufficient condition (18) is not satisfied, then at least one firm offers a contract such that the effort level set for one type θ worker is distorted.

Let us start with the case in which $k_F > k_N$, and let us check how the rent-extraction, efficiency trade-off faced by the two firms affects the two terms of inequality (21).

When only UIC_N is binding, the non-profit firm is bound to pay an information rent to the mimickers, i.e. to low-ability workers, so that $U_N^*\left(\overline{\theta}\right) > U_N^B\left(\overline{\theta}\right)$. Moreover, firm N also distorts the effort required from the high-ability type $\underline{\theta}$ upwards to save in informations rents left to type $\overline{\theta}$, whereby $x_N^*\left(\underline{\theta}\right) > x_N^{FB}\left(\underline{\theta}\right)$. And since the utilities that a firm leaves to different types of agents move in opposite directions with respect to the benchmark, it also holds that $U_N^*\left(\underline{\theta}\right) < U_N^B\left(\underline{\theta}\right)$(see Appendix A.3). Consider now the for-profit firm. For every θ, it still sets $U_F^*\left(\theta\right)$ according to its reaction function, nevertheless, being utilities strategic complements, $U_F^*\left(\overline{\theta}\right)$ increases as a consequence of the increase in $U_N^*\left(\overline{\theta}\right)$, and $U_F^*\left(\underline{\theta}\right)$ decreases as a consequence of the decrease in $U_N^*\left(\underline{\theta}\right)$, but these effects are of second order. It then follows that the wages of low-ability workers increase with respect to the benchmark, i.e. $w_i^*\left(\overline{\theta}\right) > w_i^B\left(\overline{\theta}\right)$ for each $i = F, N$, whereas the wage offered by firm F to high-ability workers decreases with respect to the benchmark, i.e. $w_F^*\left(\underline{\theta}\right) < w_F^B\left(\underline{\theta}\right)$, and the effect on the wage offered by firm N to high-ability workers is ambiguous, i.e. $w_N^*\left(\underline{\theta}\right) \gtrless w_N^B\left(\underline{\theta}\right)$. Moreover, for low-ability workers, the labor donative effect, i.e. the first term in equation (21), is lower with respect to the benchmark, so that $\widehat{\gamma}^*\left(\overline{\theta}\right) < \widehat{\gamma}^B\left(\overline{\theta}\right)$, and the term related to the selection effect does not change because there are no allocative distortions for types $\overline{\theta}$. Therefore, low-ability workers still experience a non-profit wage penalty, which is nonetheless reduced with respect to the benchmark. As for high-ability workers, they also experience a wage penalty, but it is ambiguous whether it is lower at the incentive than at the benchmark contracts. This might well be be the case given that $w_F^*\left(\underline{\theta}\right) < w_F^B\left(\underline{\theta}\right)$.

Let us then move to consider the instance in which both UIC_N and DIC_F are binding. The main difference with respect to the preceding case is that now $x_F^*\left(\overline{\theta}\right)$ is distorted downward. This reinforces the effect on the non-profit wage penalty for low-ability workers, because the selection effect is positive but smaller compared to the benchmark. So the wage penalty is reduced with respect to the benchmark contracts for low-ability workers. As for high-ability workers it is ambiguous whether the wage penalty is higher or lower with respect to the benchmark, but it seems more likely that it be higher, given that $w_F^*\left(\underline{\theta}\right) > w_F^B\left(\underline{\theta}\right)$ while the effect on $w_N\left(\underline{\theta}\right)$ is uncertain.

Conclusions which are symmetric to the ones above can be drawn for the case of positive selection and $k_N > k_F$.

The proposition that follows provides a synthesis of our results.

Proposition 10 ***Wage differentials at the incentive contracts.*** *(a) If $k_F > k_N$, a wage differential penalizing all non-profit workers exists and it is always lower than at the benchmark contracts for low-ability workers. (b) If $k_N > k_F$, both wage penalties and wage premia for non-profit workers might be observed. For low-ability workers, if a non-profit wage premium is in place, it is lower than at the* *benchmark contracts whereas if a non-profit wage penalty is in place, it is higher than at the benchmark contracts.*

To sum up, we can conclude that, in general: (i) the revenue appropriation constraint of the non-profit firm, i.e. $\alpha_N < 1$, and (ii) labor donations from employees, captured by $\hat{\gamma}^*(\theta)$, push towards a wage penalty for workers employed at the non-profit organization. Nonetheless, our model also predicts that wage premia for non-profit employees are possible, but only when the non-profit organization benefits from a competitive advantage and this advantage is sufficiently high. In such a case, the non-profit firm benefits most from hiring high-ability workers and, despite labor donations received from all its employees, it pays both high- and low-ability employees a compensation that is larger than the one offered by the for-profit firm.

Finally, the allocative distortions introduced by the incentive contracts are such that, for low-ability workers, the difference in effort levels required by the two firms shrinks with respect to the benchmark. Then the selection effects (see the second term of equation 21) for low-productivity workers are reduced and this is the reason why the wage differential for low-ability workers is smaller than at the benchmark. This is not necessarily true for high-ability employees, because the difference in effort levels required by the two firms can either increase of decrease relative to the benchmark according to which firm holds the competitive advantage.

6 Concluding remarks

How does asymmetric information in the labor market affect the competition between for-profit and non-profit organizations willing to attract the most talented and motivated workers? In our model workers are willing to donate a part of their labor to the non-profit firm because the latter is committed to a mission and sacrifices some revenues in the social interest. While, at the equilibrium, workers with high motivation are always hired by the non-profit firm, one of the two organizations, i.e. the one which is characterized by a competitive advantage, succeeds in attracting the largest share of high ability workers. This selection pattern is confirmed and exacerbated under asymmetric information when screening contracts are analyzed. For example, in the likely case in which committing to the non-profit status makes the non-profit firm the weaker competitor ($k_N < k_F$), the latter is able to hire only a lower share of talented workers and attracts instead a larger share of low-ability workers relative to the benchmark.

As for wage differentials, our model allows us to decompose the gap in the wages offered by the two firms in two terms: the first term quantifies labor donations and pushes towards a wage penalty for employees at the non-profit firm. The second term describes instead the selection effects: it can be positive or negative according to the selection pattern of ability into the non-profit firm. Both terms are affected by possible distortions due to asymmetric information. Our model is sufficiently rich to account for both non-profit wage penalties and wage premia, being thus coherent with the mixed empirical evidence.

The model's results crucially depend on the relative magnitude of the two firms' marginal revenues which are summarized by the parameters $k_N = p_N \alpha_N c_N$ and $k_F = p_F c_F$. Here we would like to consider the variables defining the two parameters (c_i and p_i in particular) and provide some statistics and real world examples. In particular, we argue that the case $k_N > k_F$ is empirically relevant, either because $c_N > c_F$ or because $p_N > p_F$ or the two together, despite the non-profit revenue appropriation constraints.[33]

The parameter c_i denotes the marginal product of labor and captures the productive efficiency of a firm. Whether for-profit or non-profit competitors are more efficient is an empirical question and the evidence on this matter is mixed. As an example of the variability of empirical results, we report some evidence on productivity of for-profit and non-profit hospitals reviewed Barros and Siciliani (2012). Evidence reported in Sherman *et al.* (1997) overall suggests that non-profit hospitals in the U.S. are not very different in economic efficiency from for-profit hospitals. In the same way, Farsi and Filippini (2008) find no significant differences by ownership in Swiss hospitals. In our model this is coherent with the case in which $c_N = c_F$, namely the instance in which the two firms are endowed with the same technology. More recently, Shen et al. (2007) find that for-profit hospitals in the U.S. tend to be more efficient than non-profit ones. In our model, this corresponds to $c_N < c_F$. Finally, differently from the previously mentioned papers, Rosko (2001) analyzes a sample of 1,631 hospitals in the U.S. during the period 1990-1996 and finds that non-profit hospitals are more productive. This may correspond to a case where $c_N > c_F$.

As a second example about heterogeneity in firms' productivity, we report some figures about for-profit and non-profit colleges in the U.S. As the 2012 U.S. Senate Committee on Health, Labor, Education and Pensions indicates, on average, for-profits spent \$3,017 per student on instructional costs *vs* \$15,321 at private non-profit colleges. Average tuition cost at for-profit colleges is \$31,000 after grants *vs* \$26,600 for non-profit colleges. Importantly, 28% of for-profit college students graduate with a four-year degree *vs* 65% at private, non-profit colleges. Finally, for-profit schools spent \$8 per student on research *vs* \$5,887 per student at private non-profits. Overall, the previous statistics motivate the title of the report "For-Profit Higher Education: The Failure to Safeguard the Federal Investment and Ensure Student Success" and suggest that, on average, U.S. for-profit colleges *are not* more efficient than non-profit ones. Again, this might correspond to a case where $c_N > c_F$.

Let us turn now to possible differences in prices. We have already provided some examples in the model set up. Here, we would like to emphasize that the case $p_N > p_F$ has empirical relevance. Consider

[33] Recall that the inequality $k_N > k_F$ implies positive selection of ability for the non-profit firm and it is also a necessary condition for non-profit wage premia.

consumption prices observed in markets where standard and mission-oriented/ethical firms coexist. Surveys of ethical shopping show that many (ethical) consumers are willing to pay higher prices for goods and service produced by do-gooders companies. Moreover, ethical consumers have a relatively high income, education, and social status. But what about price differences between standard companies and mission-oriented firms for comparable products? A research conducted by the U.K. magazine Ethical Consumer shows that ethical products have higher prices.[34]

As a last remark, with respect to the debate about the desirability of a wage cap for CEOs of non-profit firms, our model shows that competition between for-profit and non-profit firms is beneficial since it allows to decrease allocative distortions in the screening contracts. Indeed, when the two firms are sufficiently similar to each other, then equilibrium allocations are efficient. A wage cap would obviously interfere with such market forces and would impair efficiency.

More generally, our model suggests that the revenue constraint characterizing non-profit firms should be sufficient enough to generate labor donations from motivated workers but should go in the direction of making the possible differences in marginal revenues between for-profit and non-profit firms vanish. This would allow competition to restore efficiency.

References

[1] Armstrong M. (1996), "Multiproduct Nonlinear Pricing", *Econometrica*, 64, 51-75.

[2] Armstrong M. and J. Vickers (2001), "Competitive Price Discrimination", *The RAND Journal of Economics,* 32(4), 1-27.

[3] Auriol, E. and S. Brilon (2014): "Anti-social behavior in profit and nonprofit organizations", *Journal of Public Economics* 117, 149-161.

[4] Bargain, O. and B. Melly (2008): "Public Sector Pay Gap in France: New Evidence Using Panel Data", IZA Discussion Paper n. 3427.

[5] Barigozzi F. and N. Burani (2016a), "Screening Workers for Ability and Motivation", *Oxford Economic Papers*, 68(2), 627-650.

[6] Barigozzi F. and N. Burani (2016b), "Competition and Screening with Motivated Health Professionals", *Journal of Health Economics*, forthcoming, doi: 10.1016/j.jhealeco.2016.06.003

[34] For example, the article reports that "Greener electricity tariffs certainly do carry a premium, on average £169,67 more expensive per year". Moreover, "In the household cleaners...ethical brands are coming out slightly more expensive" and also "In the cosmetic section the ethical brands were considerably more expensive". See "The price of ethics" in http://www.ethicalconsumer.org.

[7] Barigozzi F., N. Burani and D. Raggi (2016), "Productivity Crowding-Out in Labor Markets with Motivated Workers", working paper n. 883, Department of Economics, University of Bologna.

[8] Barigozzi F. and G. Turati (2012), "Human Health Care and Selection Effects. Understanding Labour Supply in the Market for Nursing", *Health Economics* 21(4), 477-483.

[9] Barros P.P. and L. Siciliani (2012), "*Public and Private Sector Interface*". In: Pauly M, McGuire T, Barros P, editors. Handbook of Health Economics. Volume 2. Elsevier; 2012. pp. 927-1002.

[10] Basov S. (2005), *Multidimensional Screening*, Springer-Verlag, Heidelberg.

[11] Bénabou R. and J. Tirole (2010), "Individual and Corporate Social Responsibility", *Economica,* 77, 1-19.

[12] Bénabou R. and J. Tirole (2016), "Bonus Culture: Competitive Pay, Screening, and Multitasking", *Journal of Political Economy*, 124(2), 305-370.

[13] Besley T. and M. Ghatak (2005), "Competition and Incentives with Motivated Agents", *American Economic Review* 95(3), 616-636.

[14] Besley, T. and Ghatak, M. (2007) "Retailing public goods: the economics of corporate social responsibility", *Journal of Public Economics* 91, 1645-1663.

[15] Biglaiser G. and C. Mezzetti (2000), "Incentive Auctions and Information Revelation", *The RAND Journal of Economics* 31(1), 145-164.

[16] Borjas, G.J., H.E. III Frech and P.B. Ginsburg (1983), "Property Rights and Wages: The Case of Nursing Homes", *Journal of Human Resources*, 18, 231-246.

[17] Delfgaauw J. and R. Dur (2007), "Signaling and screening of workers' motivation", *Journal of Economic Behavior and Organization* 62, 605-624.

[18] Delfgaauw J. and R. Dur (2008), "Incentives and workers' motivation in the public sector", *The Economic Journal* 118, 171-191.

[19] Delfgaauw J. and R. Dur (2010), "Managerial talent, motivation, and self-selection into public management", *Journal of Public Economics* 94(9-10), 654-660.

[20] DeVaro J., N. Maxwell and H. Morita (2015), "Compensation and Intrinsic Motivation in Nonprofit and For-Profit Organizations", Center for Economic Institutions, Hitotsubashi University, Working Paper Series No. 2015-10.

[21] Disney R. and A. Gosling (1998), "Does it Pay to Work in the Public Sector?" *Fiscal Studies* 19, 347-374.

[22] Farsi, M. and Filippini, M. (2008), "Effects of ownership, subsidization and teaching activities on hospital costs in Switzerland", *Health Economics*, 17(3), 335-350.

[23] Francois P. (2000), “Public Service Motivation as an Argument for Government Provision”, *Journal of Public Economics* 78(3), 275-299.

[24] Francois P. (2003), “Not-for-profit Provision of Public Services”, *Economic Journal* 113(486), C53-C61.

[25] Ghatak M. and H. Mueller (2011), “Thanks for nothing? Not-for-profits and motivated agents”, *Journal of Public Economics* 95, 94-105.

[26] Gibbons R. and L. Katz (1992), “Does Unmeasured Ability Explain Inter-Industry Wage Differentials?”, *Review of Economic Studies* 59, 515-535.

[27] Glaeser E.L., A. Shleifer (2001), “Not-for-profit entrepreneurs”, *Journal of Public Economics* 81, 99-115.

[28] Goddeeris J.H. (1988), “Compensating Differentials and Self-Selection: An Application to Lawyers”, *Journal of Political Economy* 96(2), 411-428.

[29] Gregg, P., P. Grout, A. Ratcliffe, S. Smith, and F. Windmeijer, (2011), “How important is pro-social behaviour in the delivery of public services?”, Journal of Public Economics 95(7-8), 758-766.

[30] Handy F. and E. Katz (1998), “The wage differential between nonprofit institutions and corporations: getting more by paying less?”, *Journal of Comparative Economics* 26, 246-261.

[31] Hansmann H. (1996), *The Ownership of Enterprise*, Harvard University Press, Cambridge MA.

[32] Heyes A. (2005) “The economics of vocation *or* -why is a badly paid nurse a good nurse-?”, *Journal of Health Economics* 24, 561-569.

[33] Hwang H., W.R. Reed and C. Hubbard (1992), “Compensating Wage Differentials and Unobserved Productivity”, *Journal of Political Economy* 100(4), 835-858.

[34] James P. (2002), “Do Non-profit Enterprises Pay More For Their Labor Inputs? An Examination of Hospital Behavior”. Working paper, Stern School of Business. New York University.

[35] Jones D.B. (2015), “The supply and demand of motivated labor: When should we expect to see nonprofit wage gaps?”, *Labour Economics*, 32, 1-14.

[36] Kitzmueller and Shimshack (2012), “Economic Perspectives on Corporate Social Responsibility”, *Journal of Economic Literature*, 50(1), 51-84.

[37] Kosfeld M. and F.A. von Siemens (2009), “Worker Self-Selection and the Profits from Cooperation”, *Journal of the European Economic Association*, 7(2-3), 573-582.

[38] Kosfeld M. and F.A. von Siemens (2011), “Competition, Cooperation, and Corporate Culture”, *RAND Journal of Economics*, 42(1), 23-43.

[39] Leete L. (2001), "Whither the Nonprofit Wage Differential? Estimates from the 1990 Census", *Journal of Labor Economics*, 19(1), 136-170.

[40] Laffont J.-J. and D. Martimort (2002), *The theory of Incentives: The Principal-Agent Model*, Princeton University Press.

[41] Laffont J-J., E. Maskin and J.-C. Rochet (1987), "Optimal Non-Linear Pricing with Two-dimensional Characteristics" in T. Groves, R. Radner and S. Reiter, eds., *Information, Incentives and Economic Mechanisms*, University of Minnesota Press, Minneapolis, 256-266.

[42] Lehmann E., L. Simula and A. Tannoy (2014), "Tax Me if You Can! Optimal Nonlinear Income Tax Between Competing Governments", *Quarterly Journal of Economics*, 129(4), 1995-2030.

[43] Makris M. (2009), "Incentive for Motivated Agents Under an Administrative Constraint", *Journal of Economic Behavior and Organization*, 71, 428-440.

[44] Makris, M. and L. Siciliani (2013), "Optimal Incentive Schemes for Altruistic Providers", *Journal of Public Economic Theory*, 15(5), 675-699.

[45] Martimort, D. (1992), "Multi-principaux avec anti-selection", *Annales d'Economie et de Statistique*, 28, 1-37.

[46] Melly, B. (2005), "Public–private wage differentials in Germany: evidence from quantile regression", *Empirical Economics* 30(2), 505-520.

[47] Mocan N. and E. Tekin (2003), "Nonprofit sector and part-time work: an analysis of employer–employee matched data of child care workers", *Review of Economics and Statistics* 85(1), 38-50.

[48] Murdock K. (2002), "Intrinsic Motivation and Optimal Incentive Contracts", *RAND Journal of Economics,* 33(1), 650-671.

[49] Prendergast C. (2007), "The Motivation and Bias of Bureaucrats", *American Economic Review* 97(1), 180-196.

[50] Preston A. (1988). "The Effects of Property Rights on Labor Costs of Nonprofit Firms: An Application to the Day Care Industry." *Journal of Industrial Economics*, 36, 337-350.

[51] Preston, A. (1989), "The nonprofit worker in a for-profit world", *Journal of Labor Economics* 7, 438-463.

[52] Rochet J.-C. and P. Chonè (1998), "Ironing, Sweeping, and Multidimensional Screening" *Econometrica*, 66, 783-826.

[53] Rochet J.-C. and L.A. Stole (2002), "Nonlinear Pricing with Random Participation" *Review of Economic Studies*, 69, 277-311.

[54] Roomkin M. and B. Weisbrod (1999): "Managerial Compensation and Incentives in For-Profit and Nonprofit firms", *Journal of Law, Economics and Organization*, 15(3), 750-781.

[55] Rosko, M. D. (2001), "Cost efficiency of US hospitals: A stochastic frontier approach", *Health Economics*, 10(6), 539-551.

[56] Rosen, S. (1986), "The Theory of Equalizing Differences", in: Ashenfelter, O., Layard, R. (Eds.), *Handbook of Labor Economics*, vol. 1, North Holland, 641-692.

[57] Scherman F., Goodman A.C. and M. Stano (1997), *The Economics of Health and Health Care*, second edition, chapter 18, Prentice Hall.

[58] Shen, Y.C., K. Eggleston, J. Lau and C. H. Schmid (2007), "Hospital ownership and financial performance: What explains the different empirical literature findings?" *Inquiry* 44, 41-68.

[59] Stole, L.A. (1992), "Mechanism Design Under Common Agency", mimeo, University of Chicago.

[60] Stole L.A. (1995), "Nonlinear pricing and oligopoly", *Journal of Economics & Management Strategy* 4, 529-562.

A Appendix

A.1 Proof of Lemmata 1 and 2

Let us go back to the equilibrium utilities of the benchmark case in Section 3 and let us write them as follows

$$U_i^B(\theta) = w_i^B(\theta) - \frac{1}{2}\theta\left(x_i^B(\theta)\right)^2$$

with $\theta \in \{\underline{\theta}, \overline{\theta}\}$ and $i = F, N$. Furthermore, let us define the function $\phi_i^B\left(\underline{\theta}, \overline{\theta}\right)$ as the difference between the utility that type $\underline{\theta}$ receives from firm i when revealing her true type and the utility that type $\underline{\theta}$ would receive from the same firm i when claiming that her type is $\overline{\theta}$, if exerting the first-best level of effort and receiving a compensation as in the benchmark contract. Thus, function $\phi_i^B\left(\underline{\theta}, \overline{\theta}\right)$ corresponds to the downward incentive constraint DIC whereby type $\underline{\theta}$ is not attracted by the contract that firm i offers to type $\overline{\theta}$, conditional on firm i requiring all agents to exert first-best effort levels and giving compensation schemes as in the benchmark contracts. When $\phi_i^B\left(\underline{\theta}, \overline{\theta}\right) > 0$ it means that DIC_i is always slack at the benchmark contract, or else that the benchmark contract is downward incentive compatible for firm $i = F, N$. Then

$$\phi_i^B\left(\underline{\theta}, \overline{\theta}\right) = w_i^B(\underline{\theta}) - \frac{1}{2}\underline{\theta}\left(x_i^B(\underline{\theta})\right)^2 - \left(w_i^B(\overline{\theta}) - \frac{1}{2}\underline{\theta}\left(x_i^B(\overline{\theta})\right)^2\right).$$

Notice that, in the above expression, the consequence of type $\underline{\theta}$ mimicking type $\overline{\theta}$ is visible directly in the cost of effort. All other effects are mediated by type $\underline{\theta}$ choosing effort $x_i^B\left(\overline{\theta}\right)$ instead of effort $x_i^B(\underline{\theta})$. Likewise, one can obtain functions $\phi_i^B\left(\overline{\theta}, \underline{\theta}\right)$ for each firm $i = F, N$, reverting the roles of the ability types. When $\phi_i^B\left(\overline{\theta},\underline{\theta}\right)>0$, UIC_i is always slack, again conditional on firm i requiring all agents to exert first-best effort levels and giving compensation schemes as in the benchmark contracts. Let us rewrite functions ϕ_i^B extensively and rearrange terms. For firm N, one has

$$\phi_N^B\left(\underline{\theta},\overline{\theta}\right)=\frac{\left(2k_N^2\overline{\theta}-3k_N^2\underline{\theta}+k_F^2\overline{\theta}\right)\left(\overline{\theta}-\underline{\theta}\right)}{6\overline{\theta}\underline{\theta}},$$

where $\phi_N^B\left(\underline{\theta},\overline{\theta}\right)>0$ holds for $k_F\geq k_N$, whereas, for $k_F<k_N$, $\phi_N^B\left(\underline{\theta},\overline{\theta}\right)>0$ is satisfied if and only if

$$\frac{k_N^2-k_F^2}{2k_N^2+k_F^2}<\frac{\overline{\theta}-\underline{\theta}}{\underline{\theta}}. \tag{22}$$

Moreover,

$$\phi_N^B\left(\overline{\theta},\underline{\theta}\right)=\frac{\left(3k_N^2\overline{\theta}-2k_N^2\underline{\theta}-k_F^2\underline{\theta}\right)\left(\overline{\theta}-\underline{\theta}\right)}{6\overline{\theta}\underline{\theta}}$$

where $\phi_N^B\left(\overline{\theta},\underline{\theta}\right)>0$ holds for $k_F\leq k_N$, whereas, for $k_F>k_N$, $\phi_N^B\left(\overline{\theta},\underline{\theta}\right)>0$ is satisfied if and only if

$$\frac{k_F^2-k_N^2}{3k_N^2}<\frac{\overline{\theta}-\underline{\theta}}{\underline{\theta}}, \tag{23}$$

which corresponds to condition (18) in the main text. Considering firm F, one has

$$\phi_F^B\left(\underline{\theta},\overline{\theta}\right)=\frac{\left(2k_F^2\overline{\theta}-3k_F^2\underline{\theta}+k_N^2\overline{\theta}\right)\left(\overline{\theta}-\underline{\theta}\right)}{6\overline{\theta}\underline{\theta}},$$

with $\phi_F^B\left(\underline{\theta},\overline{\theta}\right)>0$ that holds when $k_F\leq k_N$, whereas, for $k_F>k_N$, $\phi_F^B\left(\underline{\theta},\overline{\theta}\right)>0$ is satisfied if and only if

$$\frac{k_F^2-k_N^2}{2k_F^2+k_N^2}<\frac{\overline{\theta}-\underline{\theta}}{\underline{\theta}}. \tag{24}$$

Finally,

$$\phi_F^B\left(\overline{\theta},\underline{\theta}\right)=\frac{\left(3k_F^2\overline{\theta}-k_N^2\underline{\theta}-2k_F^2\underline{\theta}\right)\left(\overline{\theta}-\underline{\theta}\right)}{6\overline{\theta}\underline{\theta}},$$

where $\phi_F^B\left(\overline{\theta},\underline{\theta}\right)>0$ is always true when $k_F\geq k_N$, whereas, for $k_F<k_N$, $\phi_F^B\left(\overline{\theta},\underline{\theta}\right)>0$ is satisfied if and only if

$$\frac{k_N^2-k_F^2}{3k_F^2}<\frac{\overline{\theta}-\underline{\theta}}{\underline{\theta}}, \tag{25}$$

which, again, corresponds to condition (18) in the main text.

Summing up, suppose that $k_F=k_N$. Then all ϕ_i^B are strictly positive and we can conclude that the benchmark contracts are incentive compatible, so that they are the optimal contracts not only when ability is observable, but also when ability is the agent's private information. Alternatively, suppose that $k_F>k_N$. Then $\phi_N^B\left(\underline{\theta},\overline{\theta}\right)$ and $\phi_F^B\left(\overline{\theta},\underline{\theta}\right)$ are strictly positive meaning that the benchmark contracts are always downward incentive compatible for firm N and upward incentive compatible for firm F, respectively. Moreover, $\phi_N^B\left(\overline{\theta},\underline{\theta}\right)>0$ holds when condition (23) is satisfied and $\phi_F^B\left(\underline{\theta},\overline{\theta}\right)>0$ holds when condition (24) is satisfied, with

$$\frac{k_F^2-k_N^2}{2k_F^2+k_N^2}<\frac{k_F^2-k_N^2}{3k_N^2}.$$

Then, all ϕ_i^B are strictly positive and both firms' benchmark contracts are incentive compatible when condition (23) holds, whereas only ϕ_F^B are strictly positive meaning that only firm F's benchmark contracts are incentive compatible when

$$\frac{k_F^2 - k_N^2}{2k_F^2 + k_N^2} < \frac{\overline{\theta} - \underline{\theta}}{\underline{\theta}} \leq \frac{k_F^2 - k_N^2}{3k_N^2},$$

which corresponds to condition (19) in the main text. Finally, when

$$\frac{\overline{\theta} - \underline{\theta}}{\underline{\theta}} \leq \frac{k_F^2 - k_N^2}{2k_F^2 + k_N^2},$$

it means that neither firm's benchmark contracts are incentive compatible. To conclude, suppose that $k_F < k_N$. Then $\phi_N^B\left(\overline{\theta}, \underline{\theta}\right)$ and $\phi_F^B\left(\underline{\theta}, \overline{\theta}\right)$ are strictly positive implying that the benchmark contracts are always upward incentive compatible for firm N and downward incentive compatible for firm F, respectively. Moreover, $\phi_N^B\left(\underline{\theta}, \overline{\theta}\right) > 0$ holds when condition (22) is satisfied and $\phi_F^B\left(\overline{\theta}, \underline{\theta}\right) > 0$ holds when condition (25) is satisfied, with

$$\frac{k_N^2 - k_F^2}{2k_N^2 + k_F^2} < \frac{k_N^2 - k_F^2}{3k_F^2}.$$

Then, all ϕ_i^B are strictly positive and both firms' benchmark contracts are incentive compatible when condition (22) holds, whereas only ϕ_N^B are strictly positive meaning that only firm N's benchmark contracts are incentive compatible when

$$\frac{k_N^2 - k_F^2}{2k_N^2 + k_F^2} < \frac{\overline{\theta} - \underline{\theta}}{\underline{\theta}} \leq \frac{k_N^2 - k_F^2}{3k_F^2},$$

which, again, corresponds to condition (19) in the main text. Finally, when

$$\frac{\overline{\theta} - \underline{\theta}}{\underline{\theta}} \leq \frac{k_N^2 - k_F^2}{2k_N^2 + k_F^2},$$

neither firm's benchmark contracts are incentive compatible.

A.2 Proof of Lemma 3

In order to prove Lemma 3, let us first consider a preliminary step. Let us express incentive constraints in terms of profit margins on each ability type (see expression 8), whereby DIC_i becomes

$$\pi_i\left(\underline{\theta}\right) - \pi_i\left(\overline{\theta}\right) \leq S_i\left(\underline{\theta}\right) - S_i\left(\overline{\theta}\right) - \frac{1}{2}\left(\overline{\theta} - \underline{\theta}\right) x_i^2\left(\overline{\theta}\right)$$

and UIC_i takes the form

$$S_i\left(\underline{\theta}\right) - S_i\left(\overline{\theta}\right) - \frac{1}{2}\left(\overline{\theta} - \underline{\theta}\right) x_i^2\left(\underline{\theta}\right) \leq \pi_i\left(\underline{\theta}\right) - \pi_i\left(\overline{\theta}\right).$$

Result 1 *(i) If DIC_i is binding for firm $i = F, N$, then profit margins are strictly decreasing in θ and $\pi_i\left(\underline{\theta}\right) > \pi_i\left(\overline{\theta}\right)$. (ii) If UIC_i is binding for firm $i = F, N$, then profit margins are strictly increasing in θ and $\pi_i\left(\overline{\theta}\right) > \pi_i\left(\underline{\theta}\right)$. (iii) If neither DIC_i nor UIC_i is binding for either firm, then profit margins can be either decreasing or increasing in θ.*

Proof. The proof of this result follows an argument similar to the one developed by Rochet and Stole (2002). When DIC_i is binding for firm $i = F, N$, effort levels are such that $x_i\left(\overline{\theta}\right) \leq x_i^{FB}\left(\overline{\theta}\right)$ and $x_i\left(\underline{\theta}\right) = x_i^{FB}\left(\underline{\theta}\right)$; namely, the high-ability type gets the first-best while the effort of the low-ability type is downward distorted. Moreover, when DIC_i is binding, one has

$$\pi_i\left(\underline{\theta}\right) - \pi_i\left(\overline{\theta}\right) = S_i\left(\underline{\theta}\right) - S_i\left(\overline{\theta}\right) - \frac{1}{2}\left(\overline{\theta} - \underline{\theta}\right) x_i^2\left(\overline{\theta}\right).$$

The right-hand-side of the above equality is minimized when $x_i\left(\overline{\theta}\right)$ is the highest possible, that is when it equals the first-best effort level. Substituting for such effort level yields

$$\begin{array}{cc} \pi_i\left(\underline{\theta}\right) - \pi_i\left(\overline{\theta}\right) = S_i\left(\underline{\theta}\right) - S_i\left(\overline{\theta}\right) - \frac{1}{2}\left(\overline{\theta} - \underline{\theta}\right) x_i^2\left(\overline{\theta}\right) & \geq \\ \frac{k_i^2\left(\overline{\theta}-\underline{\theta}\right)}{2\underline{\theta}\overline{\theta}} - \frac{1}{2}\left(\overline{\theta} - \underline{\theta}\right)\frac{k_i^2}{\overline{\theta}} = \frac{k_i^2\left(\overline{\theta}-\underline{\theta}\right)^2}{2\overline{\theta}\underline{\theta}} & > 0 \end{array}.$$

Similarly, when UIC_i is binding for firm $i = F, N$, effort levels are such that $x_i\left(\overline{\theta}\right) = x_i^{FB}\left(\overline{\theta}\right)$ and $x_i\left(\underline{\theta}\right) \geq x_i^{FB}\left(\underline{\theta}\right)$; namely, the low-ability type gets the first-best while the effort of the low-ability type is distorted upwards. Moreover, when UIC_i is binding, one has

$$\pi_i\left(\underline{\theta}\right) - \pi_i\left(\overline{\theta}\right) = S_i\left(\underline{\theta}\right) - S_i\left(\overline{\theta}\right) - \frac{1}{2}\left(\overline{\theta} - \underline{\theta}\right) x_i^2\left(\underline{\theta}\right).$$

The right-hand-side of the above equality is maximized when $x_i\left(\underline{\theta}\right)$ is the lowest possible, that is when it equals the first-best effort level. Substituting for such effort level yields

$$\begin{array}{cc} \pi_i\left(\underline{\theta}\right) - \pi_i\left(\overline{\theta}\right) = S_i\left(\underline{\theta}\right) - S_i\left(\overline{\theta}\right) - \frac{1}{2}\left(\overline{\theta} - \underline{\theta}\right) x_i^2\left(\underline{\theta}\right) & \leq \\ \frac{k_i^2\left(\overline{\theta}-\underline{\theta}\right)}{2\underline{\theta}\overline{\theta}} - \frac{1}{2}\left(\overline{\theta} - \underline{\theta}\right)\frac{k_i^2}{\underline{\theta}} = -\frac{k_i^2\left(\overline{\theta}-\underline{\theta}\right)^2}{2\underline{\theta}\overline{\theta}} & < 0 \end{array}.$$

When neither DIC_i nor UIC_i is binding, then each firm sets all effort levels at the first-best and profit margins can be either positive or negative. ■

Let us then move to the actual proof of Lemma 3. Suppose that there is negative selection of ability for firm N and thus that $\widehat{\gamma}\left(\underline{\theta}\right) > \widehat{\gamma}\left(\overline{\theta}\right)$ holds, whereby

$$U_F\left(\underline{\theta}\right) - U_N\left(\underline{\theta}\right) > U_F\left(\overline{\theta}\right) - U_N\left(\overline{\theta}\right) \Longleftrightarrow 1 - \left(U_F\left(\overline{\theta}\right) - U_N\left(\overline{\theta}\right)\right) > 1 - \left(U_F\left(\underline{\theta}\right) - U_N\left(\underline{\theta}\right)\right).$$

Take the problem PN of the non-profit firm (see page 13) subject to DIC_N and UIC_N. Build the Lagrangian associated with this problem, where λ_N^D and λ_N^U are the multipliers associated with DIC_N and UIC_N, respectively

$$\begin{array}{rl} \mathcal{L}_N = & \nu\left(k_N x_N\left(\underline{\theta}\right) - \frac{1}{2}\underline{\theta} x_N^2\left(\underline{\theta}\right) - U_N\left(\underline{\theta}\right)\right)\left(1 - \left(U_F\left(\underline{\theta}\right) - U_N\left(\underline{\theta}\right)\right)\right) \\ & + (1-\nu)\left(k_N x_N\left(\overline{\theta}\right) - U_N\left(\overline{\theta}\right) - \frac{1}{2}\overline{\theta} x_N^2\left(\overline{\theta}\right)\right)\left(1 - \left(U_F\left(\overline{\theta}\right) - U_N\left(\overline{\theta}\right)\right)\right) \\ & + \lambda_N^D\left(U_N\left(\underline{\theta}\right) - U_N\left(\overline{\theta}\right) - \frac{1}{2}\left(\overline{\theta} - \underline{\theta}\right) x_N^2\left(\overline{\theta}\right)\right) + \lambda_N^U\left(U_N\left(\overline{\theta}\right) - U_N\left(\underline{\theta}\right) + \frac{1}{2}\left(\overline{\theta} - \underline{\theta}\right) x_N^2\left(\underline{\theta}\right)\right) \end{array}. \quad (26)$$

The first-order conditions relative to utilities are

$$\begin{aligned}\frac{\partial \mathcal{L}_N}{\partial U_N(\underline{\theta})} = & -\nu\left(1-\left(U_F(\underline{\theta})-U_N(\underline{\theta})\right)\right) \\ & +\nu\left(k_N x_N(\underline{\theta})-\tfrac{1}{2}\underline{\theta} x_N^2(\underline{\theta})-U_N(\underline{\theta})\right)+\lambda_N^D-\lambda_N^U=0\end{aligned} \tag{N1}$$

$$\begin{aligned}\frac{\partial \mathcal{L}_N}{\partial U_N(\overline{\theta})} = & -(1-\nu)\left(1-\left(U_F(\overline{\theta})-U_N(\overline{\theta})\right)\right) \\ & +(1-\nu)\left(k_N x_N(\overline{\theta})-\tfrac{1}{2}\overline{\theta} x_N^2(\overline{\theta})-U_N(\overline{\theta})\right)-\lambda_N^D+\lambda_N^U=0\end{aligned} \tag{N2}$$

Consider the following two cases.

(a) Suppose that $\lambda_N^U>0$ while $\lambda_N^D=0$. Then DIC is slack while UIC is binding. Then equations ($N1$) and ($N2$) become

$$\begin{aligned}\frac{\partial \mathcal{L}_N}{\partial U_N(\underline{\theta})} = & -\nu\left(1-\left(U_F(\underline{\theta})-U_N(\underline{\theta})\right)\right) \\ & +\nu\left(k_N x_N(\underline{\theta})-\tfrac{1}{2}\underline{\theta} x_N^2(\underline{\theta})-U_N(\underline{\theta})\right)-\lambda_N^U=0\end{aligned} \tag{N1a}$$

$$\begin{aligned}\frac{\partial \mathcal{L}_N}{\partial U_N(\overline{\theta})} = & -(1-\nu)\left(1-\left(U_F(\overline{\theta})-U_N(\overline{\theta})\right)\right) \\ & +(1-\nu)\left(k_N x_N(\overline{\theta})-\tfrac{1}{2}\overline{\theta} x_N^2(\overline{\theta})-U_N(\overline{\theta})\right)+\lambda_N^U=0\end{aligned}. \tag{N2a}$$

Solving both ($N1a$) and ($N2a$) for λ_N^U yields

$$\pi_N(\underline{\theta}) \equiv k_N x_N(\underline{\theta})-\frac{1}{2}\underline{\theta} x_N^2(\underline{\theta})-U_N(\underline{\theta})>1-\left(U_F(\underline{\theta})-U_N(\underline{\theta})\right)$$

and

$$1-\left(U_F(\overline{\theta})-U_N(\overline{\theta})\right)>k_N x_N(\overline{\theta})-\frac{1}{2}\overline{\theta} x_N^2(\overline{\theta})-U_N(\overline{\theta}) \equiv \pi_N(\overline{\theta})$$

Given that, by Result 1, profit margins are increasing in θ when UIC is binding, one has that

$$1-\left(U_F(\overline{\theta})-U_N(\overline{\theta})\right)>\pi_N(\overline{\theta})>\pi_N(\underline{\theta})>1-\left(U_F(\underline{\theta})-U_N(\underline{\theta})\right)$$

which requires negative selection of ability for firm N. In other words, our initial assumption about negative selection is compatible with UIC binding for the N firm.

(b) Conversely, assume that UIC is slack while DIC is binding whereby $\lambda_N^U=0$ while $\lambda_N^D>0$. Now, first-order conditions ($N1$) and ($N2$) specify as

$$\begin{aligned}\frac{\partial \mathcal{L}_N}{\partial U_N(\underline{\theta})} = & -\nu\left(1-\left(U_F(\underline{\theta})-U_N(\underline{\theta})\right)\right) \\ & +\nu\left(k_N x_N(\underline{\theta})-\tfrac{1}{2}\underline{\theta} x_N^2(\underline{\theta})-U_N(\underline{\theta})\right)+\lambda_N^D=0\end{aligned} \tag{N1b}$$

$$\begin{aligned}\frac{\partial \mathcal{L}_N}{\partial U_N(\overline{\theta})} = & -(1-\nu)\left(1-\left(U_F(\overline{\theta})-U_N(\overline{\theta})\right)\right) \\ & +(1-\nu)\left(k_N x_N(\overline{\theta})-\tfrac{1}{2}\overline{\theta} x_N^2(\overline{\theta})-U_N(\overline{\theta})\right)-\lambda_N^D=0\end{aligned}. \tag{N2b}$$

Solving both ($N1b$) and ($N2b$) for λ_N^D yields

$$1-\left(U_F(\underline{\theta})-U_N(\underline{\theta})\right)>k_N x_N(\underline{\theta})-\frac{1}{2}\underline{\theta} x_N^2(\underline{\theta})-U_N(\underline{\theta}) \equiv \pi_N(\underline{\theta})$$

and

$$\pi_N\left(\overline{\theta}\right) \equiv k_N x_N\left(\overline{\theta}\right) - \frac{1}{2}\overline{\theta}x_N^2\left(\overline{\theta}\right) - U_N\left(\overline{\theta}\right) > 1 - \left(U_F\left(\overline{\theta}\right) - U_N\left(\overline{\theta}\right)\right).$$

Profit margins are decreasing in θ when DIC is binding and thus

$$1 - \left(U_F\left(\underline{\theta}\right) - U_N\left(\underline{\theta}\right)\right) > \pi_N\left(\underline{\theta}\right) > \pi_N\left(\overline{\theta}\right) > 1 - \left(U_F\left(\overline{\theta}\right) - U_N\left(\overline{\theta}\right)\right)$$

contradicting the fact that there's negative selection of ability for firm N.

Considering now the problem PF of the for-profit firm yields the following result: the assumption of negative selection of ability for firm N is compatible with DIC being binding and UIC being slack for firm F (because profit margins are decreasing in θ when DIC is binding); such assumption is instead incompatible with DIC being slack and UIC being binding for firm F (because profit margins are increasing in θ when UIC is binding).

Finally, when one assumes either a positive selection of ability for firm N or ability-neutrality, the argument follows the same lines and is thus left to the reader.

A.3 Negative selection: Optimal contracts when UIC binds for the non-profit firm

Suppose that $k_F > k_N$. Consider first the problem of firm F. It corresponds to (PF) at page 2 under no additional constraints, therefore firm F solves

$$\begin{array}{rl}\max_{x_F,U_F} E\left(\pi_F\right) = & \nu\left(k_F x_F\left(\underline{\theta}\right) - \frac{1}{2}\underline{\theta}x_F^2\left(\underline{\theta}\right) - U_F\left(\underline{\theta}\right)\right)\left(U_F\left(\underline{\theta}\right) - U_N\left(\underline{\theta}\right)\right) \\ & +\left(1-\nu\right)\left(k_F x_F\left(\overline{\theta}\right) - U_F\left(\overline{\theta}\right) - \frac{1}{2}\overline{\theta}x_F^2\left(\overline{\theta}\right)\right)\left(U_F\left(\overline{\theta}\right) - U_N\left(\overline{\theta}\right)\right)\end{array}.$$

The system of first-order conditions to this problem is

$$\frac{\partial E(\pi_F)}{\partial x_F(\underline{\theta})} = \nu\left(k_F - \underline{\theta}x_F\left(\underline{\theta}\right)\right)\left(U_F\left(\underline{\theta}\right) - U_N\left(\underline{\theta}\right)\right) = 0 \ , \tag{30}$$

$$\frac{\partial E(\pi_F)}{\partial x_F(\overline{\theta})} = \left(1-\nu\right)\left(k_F - \overline{\theta}x_F\left(\overline{\theta}\right)\right)\left(U_F\left(\overline{\theta}\right) - U_N\left(\overline{\theta}\right)\right) = 0 \ , \tag{31}$$

$$\frac{\partial E(\pi_F)}{\partial U_F(\underline{\theta})} = -\nu\left(U_F\left(\underline{\theta}\right) - U_N\left(\underline{\theta}\right)\right) + \nu\left(k_F x_F\left(\underline{\theta}\right) - \frac{1}{2}\underline{\theta}x_F^2\left(\underline{\theta}\right) - U_F\left(\underline{\theta}\right)\right) = 0 \ , \tag{32}$$

and finally

$$\frac{\partial E(\pi_F)}{\partial U_F(\overline{\theta})} = -\left(1-\nu\right)\left(U_F\left(\overline{\theta}\right) - U_N\left(\overline{\theta}\right)\right) + \left(1-\nu\right)\left(k_F x_F\left(\overline{\theta}\right) - U_F\left(\overline{\theta}\right) - \frac{1}{2}\overline{\theta}x_F^2\left(\overline{\theta}\right)\right) = 0 \tag{33}$$

Conditions (30) and (31) yield first-best effort levels, whereby $x_F^*\left(\theta\right) = \frac{k_F}{\theta} = x_F^{FB}\left(\theta\right)$ for all $\theta \in \left\{\underline{\theta}, \overline{\theta}\right\}$. Conditions (32) and (33) can be rewritten substituting for optimal effort levels in order to obtain

$$U_F^*\left(\underline{\theta}\right) = \frac{1}{2}\left(\frac{k_F^2}{2\underline{\theta}} + U_N\left(\underline{\theta}\right)\right) \quad \text{and} \quad U_F^*\left(\overline{\theta}\right) = \frac{1}{2}\left(\frac{k_F^2}{2\overline{\theta}} + U_N\left(\overline{\theta}\right)\right) \ . \tag{34}$$

Consider now firm N and assume that UIC is binding while DIC is slack. Its program is (PN) and the Lagrangian associated with it is

$$\mathcal{L}_N = E(\pi_N) + \lambda_N^U \left(U_N(\overline{\theta}) - U_N(\underline{\theta}) + \frac{1}{2}(\overline{\theta} - \underline{\theta}) x_N^2(\underline{\theta}) \right)$$

with $\lambda_N^U > 0$ being the Lagrange multiplier associated with UIC and $E(\pi_N)$ being the expected profits of firm N (as in equation 26). The first-order conditions with respect to effort levels are

$$\frac{\partial \mathcal{L}_N}{\partial x_N(\underline{\theta})} = \nu(k_N - \underline{\theta} x_N(\underline{\theta}))(1 - (U_F(\underline{\theta}) - U_N(\underline{\theta}))) + \lambda_N^U (\overline{\theta} - \underline{\theta}) x_N(\underline{\theta}) = 0 \tag{35}$$

and

$$\frac{\partial \mathcal{L}_N}{\partial x_N(\overline{\theta})} = (1-\nu)\left(k_N - \overline{\theta} x_N(\overline{\theta})\right)\left(1 - \left(U_F(\overline{\theta}) - U_N(\overline{\theta})\right)\right) = 0 \tag{36}$$

where, from (36), it follows that the first-best effort level is required for low-ability types and $x_N^*(\overline{\theta}) = x_N^{FB}(\overline{\theta})$, whereas, from (35), it follows that $k_N - \underline{\theta} x_N(\underline{\theta}) < 0$ whereby

$$x_N^*(\underline{\theta}) > \frac{k_N}{\underline{\theta}} = x_N^{FB}(\underline{\theta}).$$

In particular,

$$x_N^*(\underline{\theta}) = \frac{\nu k_N (1 - (U_F(\underline{\theta}) - U_N(\underline{\theta})))}{\nu \underline{\theta}(1 - (U_F(\underline{\theta}) - U_N(\underline{\theta}))) - \lambda_N^U (\overline{\theta} - \underline{\theta})}.$$

Notice that, combining the binding UIC for firm N with the negative selection of ability for firm N, one gets

$$\frac{1}{2}(\overline{\theta} - \underline{\theta}) x_N^2(\underline{\theta}) = U_N(\underline{\theta}) - U_N(\overline{\theta}) < U_F(\underline{\theta}) - U_F(\overline{\theta}).$$

Using (34), yields

$$x_N^*(\underline{\theta}) < \frac{k_F}{\sqrt{\underline{\theta}\overline{\theta}}}$$

whereby, the following chain of inequalities, which ranks the optimal effort levels, holds

$$x_N^*(\overline{\theta}) = x_N^{FB}(\overline{\theta}) < x_N^{FB}(\underline{\theta}) < x_N^*(\underline{\theta}) < \frac{k_F}{\sqrt{\underline{\theta}\overline{\theta}}} < \frac{k_F}{\underline{\theta}} = x_F^*(\underline{\theta}) = x_F^{FB}(\underline{\theta}) \tag{37}$$

Furthermore, the first-order conditions with respect to utilities are

$$\frac{\partial \mathcal{L}_N}{\partial U_N(\underline{\theta})} = -\nu(1 - (U_F(\underline{\theta}) - U_N(\underline{\theta}))) + \nu\left(k_N x_N(\underline{\theta}) - \tfrac{1}{2}\underline{\theta} x_N^2(\underline{\theta}) - U_N(\underline{\theta})\right) - \lambda_N^U = 0 \tag{38}$$

and

$$\frac{\partial \mathcal{L}_N}{\partial U_N(\overline{\theta})} = -(1-\nu)\left(1 - \left(U_F(\overline{\theta}) - U_N(\overline{\theta})\right)\right) + (1-\nu)\left(k_N x_N(\overline{\theta}) - \tfrac{1}{2}\overline{\theta} x_N^2(\overline{\theta}) - U_N(\overline{\theta})\right) + \lambda_N^U = 0\ . \tag{39}$$

Substituting for $x_N^{FB}(\overline{\theta})$ into (39) yields

$$\frac{\lambda_N^U}{(1-\nu)} = \left(1 - \left(U_F(\overline{\theta}) - U_N(\overline{\theta})\right)\right) - \left(\frac{k_N^2}{2\overline{\theta}} - U_N(\overline{\theta})\right), \tag{40}$$

whereby

$$U_N\left(\overline{\theta}\right)=\frac{\lambda_N^U}{2\left(1-\nu\right)}+\frac{1}{2}\left(\frac{k_N^2}{2\overline{\theta}}-\left(1-U_F\left(\overline{\theta}\right)\right)\right). \tag{41}$$

The second term on the right hand side of the above expression is the same as the reaction function of firm N at the benchmark contracts. Thus, expression 41 suggests that $U_N\left(\overline{\theta}\right)$ is higher than at the benchmark contracts, being $\lambda_N^U>0$, and that it is positively related to $U_F\left(\overline{\theta}\right)$, whereby strategic complementarities still exist. Indeed, substituting for $U_F\left(\overline{\theta}\right)$ given by (34) and rearranging yields

$$U_N^*\left(\overline{\theta}\right)=\frac{2\lambda_N^U}{3\left(1-\nu\right)}+\frac{1}{3}\left(\frac{k_N^2}{\overline{\theta}}+\frac{k_F^2}{2\overline{\theta}}-2\right)>\frac{1}{3}\left(\frac{k_N^2}{\overline{\theta}}+\frac{k_F^2}{2\overline{\theta}}-2\right)=U_N^B\left(\overline{\theta}\right)$$

(see expression (16) in the main text). Considering again the reaction function of firm F given by $U_F\left(\overline{\theta}\right)$ in (34), it is easy to see that an increase in $U_N\left(\overline{\theta}\right)$ triggers an increase in $U_F\left(\overline{\theta}\right)$ but the latter is of second order with respect to the former. Hence, the difference $U_F^*\left(\overline{\theta}\right)-U_N^*\left(\overline{\theta}\right)=\widehat{\gamma}^*\left(\overline{\theta}\right)$ decreases with respect to the benchmark.

Moreover, consider (38): one can rewrite it as

$$U_N\left(\underline{\theta}\right)=\frac{1}{2}\left(S_N\left(\underline{\theta}\right)-\left(1-U_F\left(\underline{\theta}\right)\right)\right)-\frac{\lambda_N^U}{2\nu},$$

where $S_N\left(\underline{\theta}\right)=k_Nx_N\left(\underline{\theta}\right)-\frac{1}{2}\underline{\theta}x_N^2\left(\underline{\theta}\right)$, which is suggestive of the strategic complementarity between $U_N\left(\underline{\theta}\right)$ and $U_F\left(\underline{\theta}\right)$ and of the fact that $U_N\left(\underline{\theta}\right)$ decreases with respect to the benchmark contract. Indeed, substituting for $U_F\left(\overline{\theta}\right)$ given by (34) and rearranging yields

$$U_N^*\left(\underline{\theta}\right)<\frac{1}{3}\left(2S_N\left(\underline{\theta}\right)+\frac{k_F^2}{2\underline{\theta}}-2\right),$$

where $S_N\left(\underline{\theta}\right)$ is smaller than at the first-best, because $x_N^*\left(\underline{\theta}\right)>x_N^{FB}\left(\underline{\theta}\right)$. Comparing this inequality with the same condition in the benchmark case, in which $\lambda_N^U=0$ and $x_N\left(\underline{\theta}\right)=x_N^{FB}\left(\underline{\theta}\right)$, it is easy to see that $U_N^*\left(\underline{\theta}\right)$ decreases with respect to the benchmark contracts since

$$U_N^*\left(\underline{\theta}\right)<\frac{1}{3}\left(2S_N\left(\underline{\theta}\right)+\frac{k_F^2}{2\underline{\theta}}-2\right)<\frac{1}{3}\left(\frac{k_N^2}{\underline{\theta}}+\frac{k_F^2}{2\underline{\theta}}-2\right)=U_N^B\left(\underline{\theta}\right)$$

Finally, $U_F^*\left(\underline{\theta}\right)$, which is the best reply to $U_N^*\left(\underline{\theta}\right)$ as in the benchmark case, also decreases when $U_N\left(\underline{\theta}\right)$ decreases, but to a lesser extent. Therefore the difference $U_F^*\left(\underline{\theta}\right)-U_N^*\left(\underline{\theta}\right)=\widehat{\gamma}^*\left(\underline{\theta}\right)$ increases with respect to the benchmark case. In sum, the negative selection of ability into firm N is reinforced when ability is the workers' private information.

To conclude, substituting for conditions (34) and (40) into equations (35) and (38), and considering the binding UIC for firm N, yields a system of two equations in two unknowns, namely $x_N\left(\underline{\theta}\right)$ and

$U_N\left(\overline{\theta}\right)$, which is the following

$$\begin{cases} \nu\left(k_N-\underline{\theta}x_N\left(\underline{\theta}\right)\right)\left(1-\frac{k_F^2}{4\underline{\theta}}+\frac{1}{2}U_N\left(\overline{\theta}\right)+\frac{1}{4}\left(\overline{\theta}-\underline{\theta}\right)x_N^2\left(\underline{\theta}\right)\right)+ & \\ +\left(1-\nu\right)\left(\overline{\theta}-\underline{\theta}\right)x_N\left(\underline{\theta}\right)\left(1-\frac{k_F^2}{4\overline{\theta}}+\frac{3}{2}U_N\left(\overline{\theta}\right)-\frac{k_N^2}{2\overline{\theta}}\right) & =0 \\ -\nu\left(1-\frac{k_F^2}{4\underline{\theta}}+\frac{1}{2}U_N\left(\overline{\theta}\right)+\frac{1}{4}\left(\overline{\theta}-\underline{\theta}\right)x_N^2\left(\underline{\theta}\right)\right)+\nu\left(k_Nx_N\left(\underline{\theta}\right)-U_N\left(\overline{\theta}\right)-\frac{1}{2}\overline{\theta}x_N^2\left(\underline{\theta}\right)\right)+ & \\ -\left(1-\nu\right)\left(1-\frac{k_F^2}{4\overline{\theta}}+\frac{3}{2}U_N\left(\overline{\theta}\right)-\frac{k_N^2}{2\overline{\theta}}\right) & =0 \end{cases}.$$

Such a system is hard to be solved analytically, because it encompasses a third degree polynomial in $x_N\left(\underline{\theta}\right)$; nonetheless, numeric solutions are quite easy to find. As an example, consider the uniform distribution of abilities, whereby $\nu=\frac{1}{2}$, let $k_F=2$ and $k_N=1$ and assume that $\overline{\theta}=\frac{3}{2}$. Then condition (19) is satisfied and the solution is such that, for firm N, $x_N^*\left(\underline{\theta}\right)=1.089>x_N^{FB}\left(\underline{\theta}\right)=1$ and $x_N^*\left(\overline{\theta}\right)=x_N^{FB}\left(\overline{\theta}\right)=\frac{2}{3}$. Moreover, $U_N^*\left(\overline{\theta}\right)=0.017094$ and $U_N^*\left(\underline{\theta}\right)=0.313\,57$. For firm F instead $x_F^*\left(\underline{\theta}\right)=x_F^{FB}\left(\underline{\theta}\right)=2$ and $x_F^*\left(\overline{\theta}\right)=x_F^{FB}\left(\overline{\theta}\right)=\frac{4}{3}$, with $U_F^*\left(\overline{\theta}\right)=0.67521$ and $U_F^*\left(\underline{\theta}\right)=1.1568$. Then, the indifferent worker with high ability has motivation $\widehat{\gamma}^*\left(\underline{\theta}\right)=U_F\left(\underline{\theta}\right)-U_N\left(\underline{\theta}\right)=1.1568-0.31357=0.843\,23$, which is higher than that of the indifferent worker with low-ability $\widehat{\gamma}^*\left(\overline{\theta}\right)=U_F\left(\overline{\theta}\right)-U_N\left(\overline{\theta}\right)=0.67521-1.7094\times 10^{-2}=0.658\,12$, in line with negative selection of ability for firm N. Finally, wages paid by firm N are $w_N^*\left(\underline{\theta}\right)=0.906\,53$ and $w_N^*\left(\overline{\theta}\right)=0.350\,43$ whereas wages paid by firm F are given by $w_F^*\left(\underline{\theta}\right)=3.1568$ and $w_F^*\left(\overline{\theta}\right)=2.0085$ with $w_i^*\left(\underline{\theta}\right)>w_i^*\left(\overline{\theta}\right)$ for $i=N,F$ but also $w_F^*\left(\underline{\theta}\right)-w_F^*\left(\overline{\theta}\right)>w_N^*\left(\underline{\theta}\right)-w_N^*\left(\overline{\theta}\right)$. For the sake of comparison, the benchmark contracts in this case would be characterized by $U_N^B\left(\underline{\theta}\right)=\frac{1}{3}>U_N^*\left(\underline{\theta}\right)$ and $U_N^B\left(\overline{\theta}\right)=0<U_N^*\left(\overline{\theta}\right)$ for firm N and by $U_F^B\left(\underline{\theta}\right)=\frac{7}{6}=1.166\,7>U_F^*\left(\underline{\theta}\right)$ and $U_F^B\left(\overline{\theta}\right)=\frac{2}{3}<U_F^*\left(\overline{\theta}\right)$ for firm F, whereby $\widehat{\gamma}^B\left(\underline{\theta}\right)=\frac{5}{6}=0.833\,33<\widehat{\gamma}^*\left(\underline{\theta}\right)$ and $\widehat{\gamma}^B\left(\overline{\theta}\right)=\frac{2}{3}>\widehat{\gamma}^*\left(\overline{\theta}\right)$. Thus, with respect to the benchmark case, for firm N the labor supply coming from low-ability workers goes down while the labor supply coming from high-ability workers goes up. As for wages, we have $w_N^B\left(\underline{\theta}\right)=\frac{5}{6}=0.833\,33<w_N^*\left(\underline{\theta}\right)$ and $w_N^B\left(\overline{\theta}\right)=\frac{1}{3}<w_N^*\left(\overline{\theta}\right)$, whereas $w_F^B\left(\underline{\theta}\right)=\frac{19}{6}=3.166\,7>w_F^*\left(\underline{\theta}\right)$ and $w_F^B\left(\overline{\theta}\right)=2<w_F^*\left(\overline{\theta}\right)$, so that all wages increase under asymmetric information about ability except for high-ability workers employed by the for-profit firm. Finally, $w_F^*\left(\underline{\theta}\right)-w_N^*\left(\underline{\theta}\right)=3.1568-0.906\,53=2.250\,3<w_F^B\left(\underline{\theta}\right)-w_N^B\left(\underline{\theta}\right)=\frac{19}{6}-\frac{5}{6}=\frac{7}{3}=2.333\,3$ and $w_F^*\left(\overline{\theta}\right)-w_N^*\left(\overline{\theta}\right)=2.0085-0.350\,43=1.658\,1<w_F^B\left(\overline{\theta}\right)-w_N^B\left(\overline{\theta}\right)=2-\frac{1}{3}=\frac{5}{3}=1.666\,7$. So the non-profit wage penalty decreases for all types of workers with respect to the benchmark contracts.

A.4 Necessity of Assumption 1

Before moving to the next case, notice that Assumption 1 in the main text is needed because, unless the difference in ability is sufficiently low that $2\underline{\theta}>\overline{\theta}$ holds, profits for firm N from type $\underline{\theta}$ are negative. Indeed, consider

$$\pi_N\left(\underline{\theta}\right)=\left(k_Nx_N\left(\underline{\theta}\right)-\frac{1}{2}\underline{\theta}x_N^2\left(\underline{\theta}\right)-U_N\left(\underline{\theta}\right)\right)$$

and substitute for

$$U_N(\underline{\theta}) = U_N(\overline{\theta}) + \frac{1}{2}(\overline{\theta} - \underline{\theta})\, x_N^2(\underline{\theta})$$

from the binding UIC. This yields

$$\pi_N(\underline{\theta}) = \left(k_N x_N(\underline{\theta}) - \frac{1}{2}\overline{\theta} x_N^2(\underline{\theta}) - U_N(\overline{\theta})\right).$$

Since $x_N(\underline{\theta}) > x_N^{FB}(\underline{\theta})$ and total surplus is decreasing in $x_N(\underline{\theta})$, it is true that

$$\pi_N(\underline{\theta}) < \left(k_N x_N^{FB}(\underline{\theta}) - \frac{1}{2}\overline{\theta} x_N^{FB}(\underline{\theta})^2 - U_N(\overline{\theta})\right) = -\frac{(\overline{\theta} - 2\underline{\theta})\, k_N^2}{2\underline{\theta}} - U_N(\overline{\theta})$$

The right-most term is strictly negative when $\overline{\theta} \geq 2\underline{\theta}$ and hence a necessary condition for principal N to make non-negative profits from the $\underline{\theta}$ type is that $\overline{\theta} < 2\underline{\theta}$. The same conclusion, although referred to either low-ability types or to the other firm, holds for all the cases that follow.

A.5 Negative selection: Optimal contracts when UIC binds for the non-profit firm and DIC binds for the for-profit firm

Suppose that $k_F > k_N$. For firm N, UIC_N is binding while DIC_N is slack. Its program (PN), the Lagrangian associated with it and the first-order conditions are the same as in the preceding case.

Consider now the problem (PF) of firm F under the constraint that DIC_F binds. The Lagrangian associated with this problem is

$$\mathcal{L}_F = E(\pi_F) + \lambda_F^D \left(U_F(\underline{\theta}) - U_F(\overline{\theta}) - \frac{1}{2}(\overline{\theta} - \underline{\theta})\, x_F^2(\overline{\theta})\right)$$

with the following first-order conditions

$$\frac{\partial \mathcal{L}_F}{\partial x_F(\underline{\theta})} = \quad \nu\,(k_F - \underline{\theta} x_F(\underline{\theta}))\,(U_F(\underline{\theta}) - U_N(\underline{\theta})) = 0\ , \tag{42}$$

$$\frac{\partial \mathcal{L}_F}{\partial x_F(\overline{\theta})} = \quad (1-\nu)\left(k_F - \overline{\theta} x_F(\overline{\theta})\right)\left(U_F(\overline{\theta}) - U_N(\overline{\theta})\right) - \lambda_F^D (\overline{\theta} - \underline{\theta})\, x_F(\overline{\theta}) = 0\ , \tag{43}$$

$$\frac{\partial \mathcal{L}_F}{\partial U_F(\underline{\theta})} = \quad -\nu\,(U_F(\underline{\theta}) - U_N(\underline{\theta})) + \nu\left(k_F x_F(\underline{\theta}) - \tfrac{1}{2}\underline{\theta} x_F^2(\underline{\theta}) - U_F(\underline{\theta})\right) + \lambda_F^D = 0\ , \tag{44}$$

and, finally,

$$\frac{\partial \mathcal{L}_F}{\partial U_F(\overline{\theta})} = \quad -(1-\nu)\left(U_F(\overline{\theta}) - U_N(\overline{\theta})\right) + (1-\nu)\left(k_F x_F(\overline{\theta}) - \tfrac{1}{2}\overline{\theta} x_F^2(\overline{\theta}) - U_F(\overline{\theta})\right) - \lambda_F^D = 0 \tag{45}$$

From (42) and (43) one gets $x_F^*(\underline{\theta}) = \frac{k_F}{\underline{\theta}} = x_F^{FB}(\underline{\theta})$ and $x_F^*(\overline{\theta}) < x_F^{FB}(\overline{\theta})$. In particular, one could write

$$x_F^*(\overline{\theta}) = \frac{(1-\nu)\, k_F \left(U_F(\overline{\theta}) - U_N(\overline{\theta})\right)}{(1-\nu)\,\overline{\theta}\left(U_F(\overline{\theta}) - U_N(\overline{\theta})\right) + \lambda_F^D (\overline{\theta} - \underline{\theta})}.$$

Notice that, combining the two binding incentive compatibility constraints, i.e. DIC_F and UIC_N, and adding negative selection of ability for firm N, one gets

$$\frac{1}{2}(\overline{\theta} - \underline{\theta})\, x_F^2(\overline{\theta}) = U_F(\underline{\theta}) - U_F(\overline{\theta}) > U_N(\underline{\theta}) - U_N(\overline{\theta}) = \frac{1}{2}(\overline{\theta} - \underline{\theta})\, x_N^2(\underline{\theta}).$$

For firm N, the solution solves the same equations as in the preceding Section A.5, whereby $x_N^*(\underline{\theta}) > x_N^{FB}(\underline{\theta})$ and $x_N^*\left(\overline{\theta}\right) = x_N^{FB}\left(\overline{\theta}\right)$. Thus, the following chain of inequalities holds with respect to optimal effort levels

$$x_F^*(\underline{\theta}) = x_F^{FB}(\underline{\theta}) > x_F^*\left(\overline{\theta}\right) > x_N^*(\underline{\theta}) > x_N^*\left(\overline{\theta}\right) = x_N^{FB}\left(\overline{\theta}\right). \tag{46}$$

As for utilities, from (44), substituting for $x_F^{FB}(\underline{\theta})$ and solving for the Lagrange multiplier, one obtains

$$\lambda_F^D = \nu\left(\left(U_F(\underline{\theta}) - U_N(\underline{\theta})\right) - \left(\frac{k_F^2}{2\underline{\theta}} - U_F(\underline{\theta})\right)\right),$$

where, since $\lambda_F^D > 0$, it must be the case that

$$U_F(\underline{\theta}) > \frac{1}{2}\left(\frac{k_F^2}{2\underline{\theta}} + U_N(\underline{\theta})\right),$$

which hints at $U_F(\underline{\theta})$ being higher than in the benchmark case. Moreover, consider conditions (38) and (44), equate and solve them for $U_F(\underline{\theta})$, obtaining

$$U_F^*(\underline{\theta}) = \frac{2}{3}\left(\frac{\lambda_F^D}{\nu} + \frac{k_F^2}{2\underline{\theta}} + \frac{1}{2}S_N^*(\underline{\theta}) - \frac{1}{2} - \frac{\lambda_N^U}{2\nu}\right).$$

The same condition at the first-best would be

$$U_F^B(\underline{\theta}) = \frac{2}{3}\left(\frac{k_F^2}{2\underline{\theta}} + \frac{1}{2}S_N^{FB}(\underline{\theta}) - \frac{1}{2}\right) = \frac{1}{3}\left(S_N^{FB}(\underline{\theta}) + \frac{k_F^2}{\underline{\theta}} - 1\right).$$

Then $U_F^*(\underline{\theta}) > U_F^B(\underline{\theta})$ if and only if

$$\frac{\left(2\lambda_F^D - \lambda_N^U\right)}{\nu} > \left(S_N^{FB}(\underline{\theta}) - S_N^*(\underline{\theta})\right) > 0,$$

a necessary condition being that $2\lambda_F^D > \lambda_N^U$. Moreover, take conditions (39) and (45), equate and solve them for $U_F\left(\overline{\theta}\right)$, yielding

$$U_F^*\left(\overline{\theta}\right) = \frac{2}{3}\left(S_F^*\left(\overline{\theta}\right) - \frac{\lambda_F^D}{(1-\nu)} + \frac{\lambda_N^U}{2(1-\nu)} + \frac{k_N^2}{4\overline{\theta}} - \frac{1}{2}\right).$$

Comparing this information rent with the benchmark utility one gets that $U_F^*\left(\overline{\theta}\right) > U_F^B\left(\overline{\theta}\right)$ if and only if

$$\frac{\left(\lambda_N^U - 2\lambda_F^D\right)}{2(1-\nu)} > S_F^{FB}\left(\overline{\theta}\right) - S_F\left(\overline{\theta}\right) > 0,$$

a necessary condition being that $\lambda_N^U > 2\lambda_F^D$. Therefore, one can conclude that $U_F^*(\underline{\theta}) > U_F^B(\underline{\theta})$ must be true because firm F must leave an information rent to high-ability workers who can mimic low-ability ones; this fact also implies that $2\lambda_F^D > \lambda_N^U$ and that $U_F^*\left(\overline{\theta}\right) < U_F^B\left(\overline{\theta}\right)$ must also holds true.

Analyzing now the selection effects, take the analogue of condition (44) at the benchmark, i.e. with $\lambda_F^D = 0$, substitute for $x_F^{FB}(\underline{\theta})$ and solve for the first-best total surplus as

$$S_F^{FB}(\underline{\theta}) \equiv \frac{k_F^2}{2\underline{\theta}} = 2U_F^B(\underline{\theta}) - U_N^B(\underline{\theta}).$$

Substituting for $S_F^{FB}\left(\underline{\theta}\right)$ into (44), and taking into account that $\lambda_F^D > 0$ in this case, one obtains

$$2\left(U_F^*\left(\underline{\theta}\right) - U_F^B\left(\underline{\theta}\right)\right) - \left(U_N^*\left(\underline{\theta}\right) - U_N^B\left(\underline{\theta}\right)\right) > 0. \tag{47}$$

Considering condition (45) and repeating the same procedure, with the difference that $S_F\left(\overline{\theta}\right) < S_F^{FB}\left(\overline{\theta}\right)$, one gets

$$U_N^*\left(\overline{\theta}\right) - U_N^B\left(\overline{\theta}\right) - 2\left(U_F^*\left(\overline{\theta}\right) - U_F^B\left(\overline{\theta}\right)\right) > 0 \tag{48}$$

Moreover, considering the programme of firm N and applying the same reasoning to the first-order conditions (38) and (39) yields

$$U_F^*\left(\underline{\theta}\right) - U_F^B\left(\underline{\theta}\right) - 2\left(U_N^*\left(\underline{\theta}\right) - U_N^B\left(\underline{\theta}\right)\right) > 0 \tag{49}$$

and

$$2\left(U_N^*\left(\overline{\theta}\right) - U_N^B\left(\overline{\theta}\right)\right) - \left(U_F^*\left(\overline{\theta}\right) - U_F^B\left(\overline{\theta}\right)\right) > 0, \tag{50}$$

respectively. Finally, putting (47) and (49) together, and rearranging, yields

$$U_F^*\left(\underline{\theta}\right) - U_N^*\left(\underline{\theta}\right) > U_F^B\left(\underline{\theta}\right) - U_N^B\left(\underline{\theta}\right) \Leftrightarrow \widehat{\gamma}^*\left(\underline{\theta}\right) > \widehat{\gamma}^B\left(\underline{\theta}\right)$$

and similarly, putting (48) and (50) together, and rearranging, yields

$$U_F^*\left(\overline{\theta}\right) - U_N^*\left(\overline{\theta}\right) < U_F^B\left(\overline{\theta}\right) - U_N^B\left(\overline{\theta}\right) \Leftrightarrow \widehat{\gamma}^*\left(\overline{\theta}\right) < \widehat{\gamma}^B\left(\overline{\theta}\right).$$

These results prove that the negative selection effect for the non-profit firm is reinforced when there is asymmetric information about workers' ability.

Finally, the complete system of equations characterizing the simultaneous solution to both firm's programmes consists of

$$\begin{cases} -\nu\left(1 - U_F\left(\overline{\theta}\right) + U_N\left(\overline{\theta}\right) - \frac{1}{2}\left(\overline{\theta} - \underline{\theta}\right)\left(x_F^2\left(\overline{\theta}\right) - x_N^2\left(\underline{\theta}\right)\right)\right)\left(\underline{\theta}x_N\left(\underline{\theta}\right) - k_N\right) + \\ +\left(1-\nu\right)\left(1 - U_F\left(\overline{\theta}\right) + 2U_N\left(\overline{\theta}\right) - \frac{k_N^2}{2\overline{\theta}}\right)\left(\overline{\theta} - \underline{\theta}\right)x_N\left(\underline{\theta}\right) = 0 \\ -\nu\left(1 - U_F\left(\overline{\theta}\right) + U_N\left(\overline{\theta}\right) - \frac{1}{2}\left(\overline{\theta} - \underline{\theta}\right)\left(x_F^2\left(\overline{\theta}\right) - x_N^2\left(\underline{\theta}\right)\right)\right) + \\ +\nu\left(k_N x_N\left(\underline{\theta}\right) - \frac{1}{2}\overline{\theta}x_N^2\left(\underline{\theta}\right) - U_N\left(\overline{\theta}\right)\right) - \left(1-\nu\right)\left(1 - U_F\left(\overline{\theta}\right) + 2U_N\left(\overline{\theta}\right) - \frac{k_N^2}{2\overline{\theta}}\right) = 0 \\ \left(1-\nu\right)\left(U_F\left(\overline{\theta}\right) - U_N\left(\overline{\theta}\right)\right)\left(k_F - \overline{\theta}x_F\left(\overline{\theta}\right)\right) + \\ -\nu\left(\left(\overline{\theta} - \underline{\theta}\right)x_F^2\left(\overline{\theta}\right) + 2U_F\left(\overline{\theta}\right) - \frac{1}{2}\left(\overline{\theta} - \underline{\theta}\right)x_N^2\left(\underline{\theta}\right) - U_N\left(\overline{\theta}\right) - \frac{k_F^2}{2\underline{\theta}}\right)\left(\overline{\theta} - \underline{\theta}\right)x_F\left(\overline{\theta}\right) = 0 \\ -\left(1-\nu\right)\left(U_F\left(\overline{\theta}\right) - U_N\left(\overline{\theta}\right)\right) + \left(1-\nu\right)\left(k_F x_F\left(\overline{\theta}\right) - \frac{1}{2}\overline{\theta}x_F^2\left(\overline{\theta}\right) - U_F\left(\overline{\theta}\right)\right) \\ -\nu\left(\left(\overline{\theta} - \underline{\theta}\right)x_F^2\left(\overline{\theta}\right) + 2U_F\left(\overline{\theta}\right) - \frac{1}{2}\left(\overline{\theta} - \underline{\theta}\right)x_N^2\left(\underline{\theta}\right) - U_N\left(\overline{\theta}\right) - \frac{k_F^2}{2\underline{\theta}}\right) = 0 \end{cases}$$

where the relevant unknowns are $x_F\left(\overline{\theta}\right)$ and $x_N\left(\underline{\theta}\right)$ on the one hand and $U_F\left(\overline{\theta}\right)$ and $U_N\left(\overline{\theta}\right)$ on the other hand.

As an example, consider the uniform distribution of abilities, whereby $\nu = \frac{1}{2}$, let $k_F = 2$ and $k_N = 1$ and assume that $\overline{\theta} = \frac{6}{5}$. Then condition (20) is satisfied and the solution is such that, for firm N,

$x_N^*(\underline{\theta}) = 1.0932 > x_N^{FB}(\underline{\theta}) = 1$ and $x_N^*(\overline{\theta}) = x_N^{FB}(\overline{\theta}) = \frac{5}{6}$. Moreover, $U_N^*(\overline{\theta}) = 0.18877$ and $U_N^*(\underline{\theta}) = 0.30828$. For firm F, instead, $x_F^*(\underline{\theta}) = x_F^{FB}(\underline{\theta}) = 2$ and $x_F^*(\overline{\theta}) = 1.6492 < x_F^{FB}(\overline{\theta}) = \frac{5}{3}$ with $U_F^*(\overline{\theta}) = 0.90489$ and $U_F^*(\underline{\theta}) = 1.1769$. Then, the indifferent worker with high ability has motivation equal to $\widehat{\gamma}^*(\underline{\theta}) = U_F^*(\underline{\theta}) - U_N^*(\underline{\theta}) = 1.1769 - 0.30828 = 0.86862$ which is higher than that of low-ability workers $\widehat{\gamma}^*(\overline{\theta}) = U_F^*(\overline{\theta}) - U_N^*(\overline{\theta}) = 0.90489 - 0.18877 = 0.71612$, in line with negative selection of ability for firm N. Finally, wages paid by firm N are $w_N^*(\underline{\theta}) = 0.90582$ and $w_N^*(\overline{\theta}) = 0.60544$ whereas wage paid by firm F are given by $w_F^*(\underline{\theta}) = 3.1769$ and $w_F^*(\overline{\theta}) = 2.5368$ with $w_i^*(\underline{\theta}) > w_i^*(\overline{\theta})$ for $i = N, F$ but also $w_F^*(\underline{\theta}) - w_F^*(\overline{\theta}) > w_N^*(\underline{\theta}) - w_N^*(\overline{\theta})$.

Finally, let us compare these results with the benchmark contracts. In this case, $U_N^B(\underline{\theta}) = \frac{1}{3} > U_N^*(\underline{\theta}) = 0.30828$ and $U_N^B(\overline{\theta}) = \frac{1}{6} < U_N^*(\overline{\theta}) = 0.18877$, moreover $U_F^B(\underline{\theta}) = \frac{7}{6} = 1.1667 < U_F^*(\underline{\theta}) = 1.1769$ and $U_F^B(\overline{\theta}) = \frac{11}{12} = 0.91667 > U_F^*(\overline{\theta}) = 0.90489$. Thus, $\widehat{\gamma}^B(\underline{\theta}) = U_F^B(\underline{\theta}) - U_N^B(\underline{\theta}) = \frac{7}{6} - \frac{1}{3} = \frac{5}{6} = 0.83333 < \widehat{\gamma}^*(\underline{\theta})$ whereas $\widehat{\gamma}^B(\overline{\theta}) = U_F^B(\overline{\theta}) - U_N^B(\overline{\theta}) = \frac{11}{12} - \frac{1}{6} = \frac{3}{4} = 0.75 > \widehat{\gamma}^*(\overline{\theta})$ so that, for firm N, the labor supply from high-ability workers decreases while the labor supply from low-ability workers increases with asymmetric information about skills. As for wages, we have $w_N^B(\underline{\theta}) = \frac{5}{6} = 0.83333 < w_N^*(\underline{\theta})$ and $w_N^B(\overline{\theta}) = \frac{7}{12} = 0.58333 < w_N^*(\overline{\theta})$, whereas $w_F^B(\underline{\theta}) = \frac{19}{6} = 3.1667 < w_F^*(\underline{\theta})$ and $w_F^B(\overline{\theta}) = \frac{31}{12} = 2.5833 > w_F^*(\overline{\theta})$, so that all wages increase under asymmetric information about ability except for low-ability workers employed by the for-profit firm. Finally, $w_F^*(\underline{\theta}) - w_N^*(\underline{\theta}) = 3.1769 - 0.90582 = 2.2711 < w_F^B(\underline{\theta}) - w_N^B(\underline{\theta}) = \frac{19}{6} - \frac{5}{6} = \frac{7}{3} = 2.3333$ and $w_F^*(\overline{\theta}) - w_N^*(\overline{\theta}) = 2.5368 - 0.60544 = 1.9314 < w_F^B(\overline{\theta}) - w_N^B(\overline{\theta}) = \frac{31}{12} - \frac{7}{12} = 2$ So the non-profit wage penalty decreases for all types of workers with respect to the benchmark contracts.

A.6 Positive selection: Optimal contracts when UIC binds for the for-profit firm

Assume that $k_F < k_N$. Consider firm F and assume that UIC is binding while DIC is slack. Its program is (PF) subject to DIC and the Lagrangian associated with it is

$$\mathcal{L}_F = E(\pi_F) + \lambda_F^U \left(U_F(\overline{\theta}) - U_F(\underline{\theta}) + \frac{1}{2}(\overline{\theta} - \underline{\theta}) x_F^2(\underline{\theta}) \right)$$

with $\lambda_F^U > 0$ being the Lagrange multiplier associated with UIC_F. The first-order conditions with respect to effort levels are

$$\begin{aligned} \frac{\partial \mathcal{L}_F}{\partial x_F(\underline{\theta})} &= \nu\left(k_F - \underline{\theta} x_F(\underline{\theta})\right)\left(U_F(\underline{\theta}) - U_N(\underline{\theta})\right) + \lambda_F^U (\overline{\theta} - \underline{\theta}) x_F(\underline{\theta}) = 0 \\ \frac{\partial \mathcal{L}_F}{\partial x_F(\overline{\theta})} &= (1-\nu)\left(k_F - \overline{\theta} x_F(\overline{\theta})\right)\left(U_F(\overline{\theta}) - U_N(\overline{\theta})\right) = 0 \end{aligned}$$

where, from the second line, one gets that the first-best effort level is required for low-ability types and $x_F(\overline{\theta}) = x_F^{FB}(\overline{\theta})$, whereas from the first line one has that

$$x_F(\underline{\theta}) > \frac{k_F}{\underline{\theta}} = x_F^{FB}(\underline{\theta}).$$

In particular,

$$x_F^*(\underline{\theta}) = \frac{\nu k_F (U_F(\underline{\theta}) - U_N(\underline{\theta}))}{\nu \underline{\theta} (U_F(\underline{\theta}) - U_N(\underline{\theta})) - \lambda_F^U (\overline{\theta} - \underline{\theta})}.$$

The first-order conditions with respect to utilities are

$$\begin{aligned}
\tfrac{\partial \mathcal{L}_F}{\partial U_F(\underline{\theta})} &= -\nu (U_F(\underline{\theta}) - U_N(\underline{\theta})) + \nu \left(k_F x_F(\underline{\theta}) - \tfrac{1}{2}\underline{\theta} x_F^2(\underline{\theta}) - U_F(\underline{\theta})\right) - \lambda_F^U = 0 \\
\tfrac{\partial \mathcal{L}_F}{\partial U_F(\overline{\theta})} &= -(1-\nu)\left(U_F(\overline{\theta}) - U_N(\overline{\theta})\right) + (1-\nu)\left(k_F x_F(\overline{\theta}) - \tfrac{1}{2}\overline{\theta} x_F^2(\overline{\theta}) - U_F(\overline{\theta})\right) + \lambda_F^U = 0 \ .
\end{aligned}$$

Substituting $x_F^{FB}(\overline{\theta})$ into the second equation yields

$$\lambda_F^U = (1-\nu)\left(U_F(\overline{\theta}) - U_N(\overline{\theta}) - \left(\frac{k_F^2}{2\overline{\theta}} - U_F(\overline{\theta})\right)\right), \tag{53}$$

whereby, because $\lambda_F^U > 0$,

$$U_F(\overline{\theta}) > \frac{1}{2}\left(\frac{k_F^2}{2\overline{\theta}} + U_N(\overline{\theta})\right). \tag{54}$$

Consider now the problem of firm N. It is the same as in the benchmark case, therefore firm N solves (PN) under no additional constraints, whereby the system of first-order conditions to this problem is

$$\begin{aligned}
\tfrac{\partial E(\pi_N)}{\partial x_N(\underline{\theta})} &= \nu (k_N - \underline{\theta} x_N(\underline{\theta}))(1 - (U_F(\underline{\theta}) - U_N(\underline{\theta}))) = 0 \\
\tfrac{\partial E(\pi_N)}{\partial x_N(\overline{\theta})} &= (1-\nu)\left(k_N - \overline{\theta} x_N(\overline{\theta})\right)\left(1 - \left(U_F(\overline{\theta}) - U_N(\overline{\theta})\right)\right) = 0 \\
\tfrac{\partial E(\pi_N)}{\partial U_N(\underline{\theta})} &= -\nu (1 - (U_F(\underline{\theta}) - U_N(\underline{\theta}))) + \nu \left(k_N x_N(\underline{\theta}) - \tfrac{1}{2}\underline{\theta} x_N^2(\underline{\theta}) - U_N(\underline{\theta})\right) = 0 \\
\tfrac{\partial E(\pi_N)}{\partial U_N(\overline{\theta})} &= -(1-\nu)\left(1 - \left(U_F(\overline{\theta}) - U_N(\overline{\theta})\right)\right) + (1-\nu)\left(k_N x_N(\overline{\theta}) - \tfrac{1}{2}\overline{\theta} x_N^2(\overline{\theta}) - U_N(\overline{\theta})\right) = 0
\end{aligned}$$

The first two conditions yield first-best effort levels, whereby $x_F^*(\theta) = \frac{k_F}{\theta} = x_F^{FB}(\theta)$ for all $\theta \in \{\underline{\theta}, \overline{\theta}\}$. The last two conditions can be rewritten substituting for optimal effort levels in order to obtain

$$U_N(\underline{\theta}) = \tfrac{1}{2}\left(\tfrac{k_N^2}{2\underline{\theta}} - 1 + U_F(\underline{\theta})\right) \quad \text{and} \quad U_N(\overline{\theta}) = \tfrac{1}{2}\left(\tfrac{k_N^2}{2\overline{\theta}} - 1 + U_F(\overline{\theta})\right) \ . \tag{55}$$

Notice that, combining the binding UIC_F with the positive selection of ability for firm N, one gets

$$\frac{1}{2}\left(\overline{\theta} - \underline{\theta}\right) x_F^2(\underline{\theta}) = U_F(\underline{\theta}) - U_F(\overline{\theta}) < U_N(\underline{\theta}) - U_N(\overline{\theta}).$$

Using (55), one gets

$$x_F(\underline{\theta}) < \frac{k_N}{\sqrt{\underline{\theta}\overline{\theta}}}$$

whereby,the following chain of inequalities holds

$$x_F^*(\overline{\theta}) = x_F^{FB}(\overline{\theta}) < x_F^{FB}(\underline{\theta}) < x_F^*(\underline{\theta}) < \frac{k_N}{\sqrt{\underline{\theta}\overline{\theta}}} < \frac{k_N}{\underline{\theta}} = x_N^*(\underline{\theta}) = x_N^{FB}(\underline{\theta}). \tag{56}$$

Notice that $x_N^*(\overline{\theta}) = x_N^{FB}(\overline{\theta})$ is missing from the above chain because its position cannot be determined unambiguously. The effort level $x_N^{FB}(\overline{\theta}) = \frac{k_N}{\overline{\theta}}$ is surely lower than $\frac{k_N}{\sqrt{\underline{\theta}\overline{\theta}}}$ and surely higher than $x_F^{FB}(\overline{\theta})$.

Moreover, $x_N^{FB}\left(\overline{\theta}\right) > x_F^{FB}\left(\underline{\theta}\right)$ if and only if $\frac{\overline{\theta}-\underline{\theta}}{\underline{\theta}} < \frac{k_N-k_F}{k_F}$ with $\frac{k_N-k_F}{k_F} > \frac{k_N^2-k_F^2}{2k_N^2+k_F^2}$ and $\frac{k_N-k_F}{k_F} > \frac{k_N^2-k_F^2}{3k_F^2}$ if and only if $2k_F > k_N > k_F$.

Analyzing utilities and following the same logic as in Appendix A.3 it is possible to show that $U_F^*\left(\overline{\theta}\right) > U_F^B\left(\overline{\theta}\right)$ which also implies that $U_N^*\left(\overline{\theta}\right) > U_N^B\left(\overline{\theta}\right)$ with $U_N^*\left(\overline{\theta}\right)$ increasing less than $U_F^*\left(\overline{\theta}\right)$ so that

$$\widehat{\gamma}^*\left(\overline{\theta}\right) = U_F^*\left(\overline{\theta}\right) - U_N^*\left(\overline{\theta}\right) > U_F^B\left(\overline{\theta}\right) - U_N^B\left(\overline{\theta}\right) = \widehat{\gamma}^B\left(\overline{\theta}\right).$$

Moreover, $U_F^*\left(\underline{\theta}\right) < U_F^B\left(\underline{\theta}\right)$ which also implies that $U_N^*\left(\underline{\theta}\right) < U_N^B\left(\underline{\theta}\right)$ but with $U_N^*\left(\underline{\theta}\right)$ decreasing less than $U_F^*\left(\underline{\theta}\right)$ whereby

$$\widehat{\gamma}^*\left(\underline{\theta}\right) = U_F^*\left(\underline{\theta}\right) - U_N^*\left(\underline{\theta}\right) < U_F^B\left(\underline{\theta}\right) - U_N^B\left(\underline{\theta}\right) = \widehat{\gamma}^B\left(\underline{\theta}\right).$$

This proves that asymmetric information about worker's ability reinforces the positive selection effect due to firm F having a competitive advantage over firm N.

Substituting for conditions (55) and (53) into equations $\frac{\partial \mathcal{L}_F}{\partial x_F(\underline{\theta})} = 0$ and $\frac{\partial \mathcal{L}_F}{\partial U_F(\underline{\theta})} = 0$, and considering UIC_F binding, yields a system of two equations in two unknowns $x_F\left(\underline{\theta}\right)$ and $U_F\left(\overline{\theta}\right)$ which is the following

$$\begin{cases} \nu\left(k_F - \underline{\theta} x_F\left(\underline{\theta}\right)\right)\left(\frac{1}{2}U_F\left(\overline{\theta}\right) + \frac{1}{4}\left(\overline{\theta}-\underline{\theta}\right)x_F^2\left(\underline{\theta}\right) - \frac{k_N^2}{4\underline{\theta}} + \frac{1}{2}\right) + \\ +\left(1-\nu\right)\left(\frac{3}{2}U_F\left(\overline{\theta}\right) - \frac{k_N^2}{4\overline{\theta}} + \frac{1}{2} - \frac{k_F^2}{2\overline{\theta}}\right)\left(\overline{\theta}-\underline{\theta}\right)x_F\left(\underline{\theta}\right) = 0 \\ -\nu\left(\frac{1}{2}U_F\left(\overline{\theta}\right) + \frac{1}{4}\left(\overline{\theta}-\underline{\theta}\right)x_F^2\left(\underline{\theta}\right) - \frac{k_N^2}{4\underline{\theta}} + \frac{1}{2}\right) + \nu\left(k_F x_F\left(\underline{\theta}\right) - U_F\left(\overline{\theta}\right) - \frac{1}{2}\overline{\theta}x_F^2\left(\underline{\theta}\right)\right) \\ -\left(1-\nu\right)\left(\frac{3}{2}U_F\left(\overline{\theta}\right) - \frac{k_N^2}{4\overline{\theta}} + \frac{1}{2} - \frac{k_F^2}{2\overline{\theta}}\right) = 0 \end{cases}$$

As an example, consider the uniform distribution of abilities, whereby $\nu = \frac{1}{2}$, let $k_F = 1$ and $k_N = \sqrt{2}$ and assume that $\overline{\theta} = \frac{5}{4}$ and $\underline{\theta} = 1$. Then condition (19) is satisfied and the solution is such that, for firm N, $x_N\left(\underline{\theta}\right) = x_N^{FB}\left(\underline{\theta}\right) = \sqrt{2} = 1.4142$ and $x_N\left(\overline{\theta}\right) = x_N^{FB}\left(\overline{\theta}\right) = \frac{4\sqrt{2}}{5} = 1.1314$. Moreover $U_N^*\left(\overline{\theta}\right) = 0.001615$ and $U_N^*\left(\underline{\theta}\right) = 0.16505$. For firm F, instead, $x_F^*\left(\underline{\theta}\right) = 1.0074 > x_F^{FB}\left(\underline{\theta}\right) = 1$ and $x_F\left(\overline{\theta}\right) = x_F^{FB}\left(\overline{\theta}\right) = \frac{4}{5} = 0.8$ with $x_N^{FB}\left(\overline{\theta}\right) > x_F^*\left(\underline{\theta}\right)$. Moreover, $U_F^*\left(\overline{\theta}\right) = 0.20323$ and $U_F^*\left(\underline{\theta}\right) = 0.33009$. Then, the motivation of the high-ability worker who is indifferent between firms is $\widehat{\gamma}^*\left(\underline{\theta}\right) = U_F^*\left(\underline{\theta}\right) - U_N^*\left(\underline{\theta}\right) = 0.33009 - 0.16505 = 0.16504$ which is lower than the motivation of the marginal worker with low-ability which is $\widehat{\gamma}^*\left(\overline{\theta}\right) = U_F^*\left(\overline{\theta}\right) - U_N^*\left(\overline{\theta}\right) = 0.20323 - 0.001615 = 0.20162$ in line with positive selection of ability for firm N. Finally wages paid by firm N firm are $w_N^*\left(\underline{\theta}\right) = 1.1651$ and $w_N^*\left(\overline{\theta}\right) = 0.80162$ whereas wages paid by firm F are given by $w_F^*\left(\underline{\theta}\right) = 0.83752$ and $w_F^*\left(\overline{\theta}\right) = 0.60323$ with $w_i\left(\underline{\theta}\right) > w_i\left(\overline{\theta}\right)$ for each $i = N, F$ and $w_N\left(\theta\right) > w_F\left(\theta\right)$ for each $\theta \in \left\{\underline{\theta}, \overline{\theta}\right\}$ but also $w_N\left(\underline{\theta}\right) - w_N\left(\overline{\theta}\right) = 1.1651 - 0.80162 = 0.36348 > w_F\left(\underline{\theta}\right) - w_F\left(\overline{\theta}\right) = 0.83752 - 0.60323 = 0.23429$. Then non-profit employees experience a wage premium for all ability levels and also higher returns to ability. The wage premium for non-profit workers arises from the difference in effort levels (firm N has a competitive advantage and thus sets higher effort levels) and it is partly offset by the compensating effect of intrinsic motivation which keeps $U_F\left(\theta\right)$ higher than $U_N\left(\theta\right)$ for all $\theta \in \left\{\underline{\theta}, \overline{\theta}\right\}$. For the sake of comparison, the benchmark contracts in this case would be characterized by $U_N^B\left(\underline{\theta}\right) = \frac{1}{6} = 0.16667 > U_N^*\left(\underline{\theta}\right)$ and $U_N^B\left(\overline{\theta}\right) = 0 < U_N^*\left(\overline{\theta}\right)$ for firm N and by

$U_F^B(\underline{\theta}) = \frac{1}{3} > U_F^*(\underline{\theta})$ and $U_F^B(\overline{\theta}) = \frac{1}{5} < U_F^*(\overline{\theta})$ for firm F, whereby $\widehat{\gamma}^B(\underline{\theta}) = U_F^B(\underline{\theta}) - U_N^B(\underline{\theta}) = \frac{1}{3} - \frac{1}{6} = 0.166\,67 > \widehat{\gamma}^*(\underline{\theta})$ and $\widehat{\gamma}^B(\overline{\theta}) = U_F^B(\overline{\theta}) - U_N^B(\overline{\theta}) = \frac{1}{5} < \widehat{\gamma}^*(\overline{\theta})$. Thus, with respect to the benchmark case, for firm N the labor supply coming from low-ability workers goes up while the labor supply coming from high-ability workers goes down. As for wages, we have $w_N^B(\underline{\theta}) = \frac{7}{6} = 1.166\,7 > w_N^*(\underline{\theta})$ and $w_N^B(\overline{\theta}) = \frac{4}{5} < w_N^*(\overline{\theta})$, whereas $w_F^B(\underline{\theta}) = \frac{5}{6} = 0.833\,33 < w_F^*(\underline{\theta})$ and $w_F^B(\overline{\theta}) = \frac{3}{5} = 0.6 < w_F^*(\overline{\theta})$, so that, with respect to the benchmark, all wages increase except for high-ability workers employed by the non-profit firm. Finally, $w_N^*(\underline{\theta}) - w_F^*(\underline{\theta}) = 1.165\,1 - 0.837\,52 = 0.327\,58 < w_N^B(\underline{\theta}) - w_F^B(\underline{\theta}) = \frac{7}{6} - \frac{5}{6} = \frac{1}{3} = 0.333\,33$ and $w_N^*(\overline{\theta}) - w_F^*(\overline{\theta}) = 0.801\,62 - 0.603\,23 = 0.198\,39 < w_N^B(\overline{\theta}) - w_F^B(\overline{\theta}) = \frac{4}{5} - \frac{3}{5} = \frac{1}{5} = 0.2$. So the non-profit wage premium decreases for all types of workers with respect to the benchmark contracts.

A.7 Positive selection: Optimal contracts when UIC binds for the for-profit firm and DIC binds for the non-profit firm

For firm F, UIC is binding while DIC is slack. Its program (PF), the Lagrangian associated with it and the first-order conditions are the same as in the preceding case.

Consider now the problem (PN) of firm N under the constraint that DIC_N binds, that is

$$U_N(\underline{\theta}) = U_N(\overline{\theta}) + \frac{1}{2}\left(\overline{\theta} - \underline{\theta}\right) x_N^2(\overline{\theta}).$$

Then the Lagrangian associated with problem (PN) is

$$\mathcal{L}_N = E(\pi_N) + \lambda_N^D \left(U_N(\underline{\theta}) - U_N(\overline{\theta}) - \frac{1}{2}\left(\overline{\theta} - \underline{\theta}\right) x_N^2(\overline{\theta}) \right)$$

with $\lambda_N^D > 0$ being the Lagrange multiplier associated with DIC_N. The first-order conditions with respect to effort levels are

$$\begin{array}{rl} \frac{\partial \mathcal{L}_N}{\partial x_N(\underline{\theta})} = & \nu\left(k_N - \underline{\theta} x_N(\underline{\theta})\right)\left(1 - \left(U_F(\underline{\theta}) - U_N(\underline{\theta})\right)\right) = 0 \\ \frac{\partial \mathcal{L}_N}{\partial x_N(\overline{\theta})} = & (1-\nu)\left(k_N - \overline{\theta} x_N(\overline{\theta})\right)\left(1 - \left(U_F(\overline{\theta}) - U_N(\overline{\theta})\right)\right) - \lambda_N^D\left(\overline{\theta} - \underline{\theta}\right) x_N(\overline{\theta}) = 0 \end{array}.$$

From the first line, one gets that the first-best effort level is required for high-ability types and $x_N^*(\underline{\theta}) = x_N^{FB}(\underline{\theta})$, whereas from the second line one has that

$$x_N^*(\overline{\theta}) < \frac{k_N}{\overline{\theta}} = x_N^{FB}(\overline{\theta});$$

In particular,

$$x_N^*(\overline{\theta}) = \frac{(1-\nu) k_N \left(1 - \left(U_F(\overline{\theta}) - U_N(\overline{\theta})\right)\right)}{(1-\nu)\overline{\theta}\left(1 - \left(U_F(\overline{\theta}) - U_N(\overline{\theta})\right)\right) + \lambda_N^D\left(\overline{\theta} - \underline{\theta}\right)}.$$

Moreover, combining the two binding incentive compatibility constraints, i.e. DIC_N and UIC_F, and adding the positive selection of ability into firm N, one gets

$$\frac{1}{2}\left(\overline{\theta} - \underline{\theta}\right) x_N^2(\overline{\theta}) = U_N(\underline{\theta}) - U_N(\overline{\theta}) > U_F(\underline{\theta}) - U_F(\overline{\theta}) = \frac{1}{2}\left(\overline{\theta} - \underline{\theta}\right) x_F^2(\underline{\theta}).$$

For firm F, the optimal allocation is such that $x_N^*(\underline{\theta}) = x_N^{FB}(\underline{\theta})$ and $x_N^*(\overline{\theta}) < x_N^{FB}(\overline{\theta})$. Thus the following chain of inequalities holds with respect to optimal effort levels

$$x_N^*(\underline{\theta}) = x_N^{FB}(\underline{\theta}) > x_N^{FB}(\overline{\theta}) > x_N^*(\overline{\theta}) > x_F^*(\underline{\theta}) > x_F^{FB}(\underline{\theta}) > x_F^*(\overline{\theta}) = x_N^{FB}(\overline{\theta}). \tag{58}$$

The first-order conditions with respect to utilities are

$$\begin{aligned}
\frac{\partial \mathcal{L}_N}{\partial U_N(\underline{\theta})} &= -\nu\left(1-\left(U_F(\underline{\theta})-U_N(\underline{\theta})\right)\right)+\nu\left(k_N x_N(\underline{\theta})-\tfrac{1}{2}\underline{\theta}x_N^2(\underline{\theta})-U_N(\underline{\theta})\right)+\lambda_N^D=0\\
\frac{\partial \mathcal{L}_N}{\partial U_N(\overline{\theta})} &= -(1-\nu)\left(1-\left(U_F(\overline{\theta})-U_N(\overline{\theta})\right)\right)+(1-\nu)\left(k_N x_N(\overline{\theta})-\tfrac{1}{2}\overline{\theta}x_N^2(\overline{\theta})-U_N(\overline{\theta})\right)-\lambda_N^D=0\ .
\end{aligned}$$

Analyzing utilities and following the same logic as in Appendix A.5 it is possible to show that $U_N^*(\underline{\theta}) > U_N^B(\underline{\theta})$ whereas $U_F^*(\underline{\theta}) < U_F^B(\underline{\theta})$ and that $U_F^*(\overline{\theta}) > U_F^B(\overline{\theta})$ whereas $U_N^*(\overline{\theta}) < U_N^B(\overline{\theta})$. Thus, it also happens that

$$\widehat{\gamma}^*(\underline{\theta}) = U_F^*(\underline{\theta}) - U_N^*(\underline{\theta}) < U_F^B(\underline{\theta}) - U_N^B(\underline{\theta}) = \widehat{\gamma}^B(\underline{\theta})$$

and that

$$\widehat{\gamma}^*(\overline{\theta}) = U_F^*(\overline{\theta}) - U_N^*(\overline{\theta}) > U_F^B(\overline{\theta}) - U_N^B(\overline{\theta}) = \widehat{\gamma}^B(\overline{\theta})$$

whereby asymmetric information about worker's ability reinforces the positive selection effect due to firm F having a competitive advantage over firm N. Finally, the system of equations to be solved is the following $\frac{\partial \mathcal{L}_N}{\partial x_N(\overline{\theta})} = 0$, $\frac{\partial \mathcal{L}_N}{\partial U_N(\overline{\theta})} = 0$ for firm N and $\frac{\partial \mathcal{L}_F}{\partial x_F(\underline{\theta})} = 0$, $\frac{\partial \mathcal{L}_F}{\partial U_F(\underline{\theta})} = 0$ for firm F. Using UIC_F and DIC_N binding, allows us to eliminate $U_F(\overline{\theta})$ and $U_N(\overline{\theta})$, respectively, from the system thus yielding

$$\begin{cases}
\nu\frac{k_N^2}{2\underline{\theta}} - \left(1-U_F(\underline{\theta})+2U_N(\underline{\theta})\right)-(1-\nu)\left(\frac{1}{2}\underline{\theta}x_N^2(\overline{\theta})-\frac{1}{2}\left(\overline{\theta}-\underline{\theta}\right)\left(x_N^2(\overline{\theta})-x_F^2(\underline{\theta})\right)-k_N x_N(\overline{\theta})\right)=0\\
(1-\nu)\left(\left(\overline{\theta}-\underline{\theta}\right)x_F^2(\underline{\theta})-\frac{1}{2}\left(\overline{\theta}-\underline{\theta}\right)x_N^2(\overline{\theta})+\frac{k_F^2}{2\overline{\theta}}\right)+\\
-\left(2U_F(\underline{\theta})-U_N(\underline{\theta})\right)+\nu\left(k_F x_F(\underline{\theta})-\frac{1}{2}\underline{\theta}x_F^2(\underline{\theta})\right)=0\\
\left(\nu\left(k_F-\underline{\theta}x_F(\underline{\theta})\right)+\left(\overline{\theta}-\underline{\theta}\right)x_F(\underline{\theta})(1-\nu)\right)\left(U_F(\underline{\theta})-U_N(\underline{\theta})\right)+\\
+\left(\overline{\theta}-\underline{\theta}\right)x_F(\underline{\theta})(1-\nu)\left(U_F(\underline{\theta})-\left(\overline{\theta}-\underline{\theta}\right)x_F^2(\underline{\theta})+\frac{1}{2}\left(\overline{\theta}-\underline{\theta}\right)x_N^2(\overline{\theta})-\frac{k_F^2}{2\overline{\theta}}\right)=0\\
(1-\nu)\left(k_N-\overline{\theta}x_N(\overline{\theta})\right)\left(1-U_F(\underline{\theta})+U_N(\underline{\theta})-\frac{1}{2}\left(\overline{\theta}-\underline{\theta}\right)\left(x_N^2(\overline{\theta})-x_F^2(\underline{\theta})\right)\right)+\\
-\nu\left(\overline{\theta}-\underline{\theta}\right)x_N(\overline{\theta})\left(1-U_F(\underline{\theta})+2U_N(\underline{\theta})-\frac{k_N^2}{2\underline{\theta}}\right)=0
\end{cases}$$

to be solved for $x_F(\underline{\theta})$ and $x_N(\overline{\theta})$ and also for $U_F(\underline{\theta})$ and $U_N(\underline{\theta})$.

As an example, consider the uniform distribution of abilities, whereby $\nu = \frac{1}{2}$, let $k_F = 1$ and $k_N = \sqrt{2}$ and assume that $\overline{\theta} = \frac{6}{5}$ and $\underline{\theta} = 1$. The solution is such that, for firm N, $x_N(\underline{\theta}) = x_N^{FB}(\underline{\theta}) = \sqrt{2} = 1.4142$ and $x_N(\overline{\theta}) = 1.1775 < x_N^{FB}(\overline{\theta}) = \frac{\sqrt{2}}{\frac{6}{5}} = 1.1785$. Moreover $U_N^*(\underline{\theta}) = 0.16653$ and $U_N^*(\overline{\theta}) = 0.027879$. For firm F, instead, $x_F^*(\underline{\theta}) = 1.0109 > x_F^{FB}(\underline{\theta}) = 1$ and $x_F^*(\overline{\theta}) = x_F^{FB}(\overline{\theta}) = \frac{5}{6} = 0.83333$ with $x_N^{FB}(\overline{\theta}) > x_F^*(\underline{\theta})$. Moreover, $U_F^*(\underline{\theta}) = 0.32886$ and $U_F^*(\overline{\theta}) = 0.22667$. Then, the motivation of the high-ability worker who is indifferent between firms is $\widehat{\gamma}^*(\underline{\theta}) = U_F^*(\underline{\theta}) - U_N^*(\underline{\theta}) = 0.32886 - 0.16653 = 0.16233$ which is lower than the motivation of the marginal worker with low-ability which is $\widehat{\gamma}^*(\overline{\theta}) =$

$U_F^*\left(\overline{\theta}\right) - U_N^*\left(\overline{\theta}\right) = 0.226\,67 - 0.027\,87\,9 = 0.198\,79$, in line with positive selection of ability for firm N. Finally wages paid by firm N firm are $w_N^*\left(\underline{\theta}\right) = 1.166\,5$ and $w_N^*\left(\overline{\theta}\right) = 0.859\,78$ whereas wages paid by firm F are given by $w_F^*\left(\underline{\theta}\right) = 0.839\,82$ and $w_F^*\left(\overline{\theta}\right) = 0.643\,34$ with $w_i^*\left(\underline{\theta}\right) > w_i^*\left(\overline{\theta}\right)$ for each $i = N, F$ and $w_N^*\left(\theta\right) > w_F^*\left(\theta\right)$ for each $\theta \in \left\{\underline{\theta}, \overline{\theta}\right\}$ but also $w_N^*\left(\underline{\theta}\right) - w_N^*\left(\overline{\theta}\right) = 1.166\,5 - 0.859\,78 = 0.306\,72 > w_F^*\left(\underline{\theta}\right) - w_F^*\left(\overline{\theta}\right) = 0.839\,82 - 0.643\,34 = 0.196\,48$. Then, as in the previous case, non-profit employees experience a wage premium for all ability levels and also higher returns to ability. The wage premium for non-profit workers arises from the difference in effort levels (firm N has a competitive advantage and thus sets higher effort levels) and it is partly offset by the compensating effect of intrinsic motivation which keeps $U_F\left(\theta\right)$ higher than $U_N\left(\theta\right)$ for all $\theta \in \left\{\underline{\theta}, \overline{\theta}\right\}$. For the sake of comparison, the benchmark contracts in this case would be characterized by $U_N^B\left(\underline{\theta}\right) = \frac{1}{6} = 0.166\,67 > U_N^*\left(\underline{\theta}\right)$ and $U_N^B\left(\overline{\theta}\right) = \frac{1}{36} = 0.027\,77\,8 < U_N^*\left(\overline{\theta}\right)$ for firm N and by $U_F^B\left(\underline{\theta}\right) = \frac{1}{3} > U_F^*\left(\underline{\theta}\right)$ and $U_F^B\left(\overline{\theta}\right) = \frac{2}{9} = 0.222\,22 < U_F^*\left(\overline{\theta}\right)$ for firm F, whereby $\widehat{\gamma}^B\left(\underline{\theta}\right) = U_F^B\left(\underline{\theta}\right) - U_N^B\left(\underline{\theta}\right) = \frac{1}{3} - \frac{1}{6} = 0.166\,67 > \widehat{\gamma}^*\left(\underline{\theta}\right)$ and $\widehat{\gamma}^B\left(\overline{\theta}\right) = U_F^B\left(\overline{\theta}\right) - U_N^B\left(\overline{\theta}\right) = \frac{2}{9} - \frac{1}{36} = \frac{7}{36} = 0.194\,44 < \widehat{\gamma}^*\left(\overline{\theta}\right)$. Thus, with respect to the benchmark case, for firm N the labor supply coming from high-ability workers goes up while the labor supply coming from low-ability workers goes down. As for wages, we have $w_N^B\left(\underline{\theta}\right) = \frac{7}{6} = 1.\,166\,7 > w_N^*\left(\underline{\theta}\right)$ and $w_N^B\left(\overline{\theta}\right) = \frac{31}{36} = 0.861\,11 > w_N^*\left(\overline{\theta}\right)$, whereas $w_F^B\left(\underline{\theta}\right) = \frac{5}{6} = 0.833\,33 < w_F^*\left(\underline{\theta}\right)$ and $w_N^B\left(\overline{\theta}\right) = \frac{23}{36} = 0.638\,89 < w_F^*\left(\overline{\theta}\right)$, so that, with respect to the benchmark, wages increase for firm F while they decrease for firm N.under asymmetric information about ability except for high-ability workers employed by the non-profit firm. Finally, $w_N^*\left(\underline{\theta}\right) - w_F^*\left(\underline{\theta}\right) = 1.166\,5 - 0.839\,82 = 0.326\,68 < w_N^B\left(\underline{\theta}\right) - w_F^B\left(\underline{\theta}\right) = \frac{7}{6} - \frac{5}{6} = \frac{1}{3} = 0.333\,33$ and $w_N^*\left(\overline{\theta}\right) - w_F^*\left(\overline{\theta}\right) = 0.859\,78 - 0.643\,34 = 0.216\,44 < w_N^B\left(\overline{\theta}\right) - w_F^B\left(\overline{\theta}\right) = \frac{31}{36} - \frac{23}{36} = \frac{2}{9} = 0.222\,22$. So the non-profit wage premium decreases for all types of workers with respect to the benchmark contracts.

Alma Mater Studiorum - Università di Bologna
DEPARTMENT OF ECONOMICS

Strada Maggiore 45
40125 Bologna - Italy
Tel. +39 051 2092604
Fax +39 051 2092664
http://www.dse.unibo.it