Supplementary Material for "Competition for Talent when Firms' Mission Matters"

by Francesca Barigozzi and Nadia Burani

1 Negative selection of ability into firm N: Optimal contracts when UIC_N binds

Suppose that $k_F > k_N$. Consider first the problem of firm F. It corresponds to programme (P_F) at page 13 of the main text under no additional incentive compatibility constraints, therefore firm F solves

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

The system of first-order conditions to this problem is

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

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

$$\frac{\partial E(\pi_F)}{\partial U_F(\underline{\theta})} = \nu\left(k_F x_F(\underline{\theta}) - \tfrac{1}{2}\underline{\theta}x_F^2(\underline{\theta}) - U_F(\underline{\theta})\right) - \nu\left(U_F(\underline{\theta}) - U_N(\underline{\theta})\right) = 0 \tag{3}$$

$$\frac{\partial E(\pi_F)}{\partial U_F(\overline{\theta})} = (1-\nu)\left(k_F x_F(\overline{\theta}) - U_F(\overline{\theta}) - \tfrac{1}{2}\overline{\theta}x_F^2(\overline{\theta})\right) - (1-\nu)\left(U_F(\overline{\theta}) - U_N(\overline{\theta})\right) = 0 \tag{4}$$

Conditions (1) and (2) yield first-best effort levels, whereby $x_F^*(\theta) = \frac{k_F}{\theta} = x_F^{FB}(\theta)$ for all $\theta \in \left\{\underline{\theta}, \overline{\theta}\right\}$. Conditions (3) and (4) can be rewritten, substituting for optimal effort levels, in order to obtain

$$U_F^*(\underline{\theta}) = \frac{1}{2}\left(\frac{k_F^2}{2\underline{\theta}} + U_N(\underline{\theta})\right) \quad \text{and} \quad U_F^*(\overline{\theta}) = \frac{1}{2}\left(\frac{k_F^2}{2\overline{\theta}} + U_N(\overline{\theta})\right). \tag{5}$$

Consider now firm N and assume that UIC_N is binding while DIC_N is slack. The programme is (P_N) and the corresponding Lagrangian is

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

with $E(\pi_N)$ being the expected profits of firm N (as in equation 26 of Appendix A.2) and $\lambda_N^U > 0$ being the Lagrange multiplier associated with UIC_N and . The first-order conditions with respect to effort levels are

$$\frac{\partial \mathcal{L}_N}{\partial x_N(\underline{\theta})} = \nu\left(k_N - \underline{\theta} x_N(\underline{\theta})\right)\left(1 - \left(U_F(\underline{\theta}) - U_N(\underline{\theta})\right)\right) + \lambda_N^U\left(\overline{\theta} - \underline{\theta}\right) x_N(\underline{\theta}) = 0 \tag{6}$$

$$\frac{\partial \mathcal{L}_N}{\partial x_N(\overline{\theta})} = (1-\nu)\left(k_N - \overline{\theta} x_N(\overline{\theta})\right)\left(1 - \left(U_F(\overline{\theta}) - U_N(\overline{\theta})\right)\right) = 0. \tag{7}$$

From (7) it follows that the first-best effort level is required for low-ability types, so $x_N^*(\overline{\theta}) = x_N^{FB}(\overline{\theta})$, whereas, from (6), 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,

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

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

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

Using (5), the expression above yields

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

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

$$x_N^*(\overline{\theta}) = x_N^{FB}(\overline{\theta}) < x_N^{FB}(\underline{\theta}) < x_N^*(\underline{\theta}) < \frac{k_F}{\sqrt{\underline{\theta}\overline{\theta}}} < \frac{k_F}{\underline{\theta}} = x_F^*(\underline{\theta}) = x_F^{FB}(\underline{\theta}). \tag{8}$$

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

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

$$\frac{\partial \mathcal{L}_N}{\partial U_N(\underline{\theta})} = \nu\left(k_N x_N(\underline{\theta}) - \tfrac{1}{2}\underline{\theta} x_N^2(\underline{\theta}) - U_N(\underline{\theta})\right) - \nu\left(1 - \left(U_F(\underline{\theta}) - U_N(\underline{\theta})\right)\right) - \lambda_N^U = 0 \tag{9}$$

$$\frac{\partial \mathcal{L}_N}{\partial U_N(\overline{\theta})} = (1-\nu)\left(k_N x_N(\overline{\theta}) - \tfrac{1}{2}\overline{\theta} x_N^2(\overline{\theta}) - U_N(\overline{\theta})\right) - (1-\nu)\left(1 - \left(U_F(\overline{\theta}) - U_N(\overline{\theta})\right)\right) + \lambda_N^U = 0 \tag{10}$$

Substituting for $x_N^{FB}(\overline{\theta})$ into (10) yields

$$\frac{\lambda_N^U}{(1-\nu)} = \left(1 - \left(U_F(\overline{\theta}) - U_N(\overline{\theta})\right)\right) - \left(\frac{k_N^2}{2\overline{\theta}} - U_N(\overline{\theta})\right), \tag{11}$$

whereby

$$U_N\left(\overline{\theta}\right)=\frac{\lambda_N^U}{2\left(1-\nu\right)}+\frac{1}{2}\left(\frac{k_N^2}{2\overline{\theta}}-\left(1-U_F\left(\overline{\theta}\right)\right)\right). \tag{12}$$

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 (12) 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 $U_F\left(\overline{\theta}\right)$, so strategic complementarities still exist. Indeed, substituting for $U_F\left(\overline{\theta}\right)$ given by (5) 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 14 in the main text). Considering again the reaction function of firm F given by $U_F\left(\overline{\theta}\right)$ in (5), 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 (9): one can rewrite it as

$$U_N\left(\underline{\theta}\right)=\frac{1}{2}\left(S_N\left(\underline{\theta}\right)-\left(1-U_F\left(\underline{\theta}\right)\right)\right)-\frac{\lambda_N^U}{2\nu}, \tag{13}$$

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 $U_N\left(\underline{\theta}\right)$ is lower than at the benchmark contract. Indeed, substituting for $U_F\left(\underline{\theta}\right)$ given by (5) 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 $U_N^*\left(\underline{\theta}\right)$ is lower than at the benchmark contracts since

$$U_N^*\left(\underline{\theta}\right)<\frac{1}{3}\left(2S_N\left(\underline{\theta}\right)+\frac{k_F^2}{2\underline{\theta}}-2\right)<\frac{1}{3}\left(\frac{k_N^2}{\underline{\theta}}+\frac{k_F^2}{2\underline{\theta}}-2\right)=U_N^B\left(\underline{\theta}\right) \tag{14}$$

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 $U_F^*\left(\underline{\theta}\right)-U_N^*\left(\underline{\theta}\right)=\widehat{\gamma}^*\left(\underline{\theta}\right)$ is bigger than in the benchmark case. In sum, the negative selection of ability into firm N is reinforced when ability is the workers' private information and screening is in place.

To conclude, substituting for conditions (5) and (11) into equations (6) and (9), and considering the binding UIC for firm N, yields a system of two equations in two unknowns, namely $x_N(\underline{\theta})$ and $U_N(\overline{\theta})$, which is the following

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

and its solution defines the optimal contracts in this case. Notice that the finding $U_i^*(\overline{\theta}) > U_i^B(\overline{\theta})$, together with inequality $\overline{\theta} < \frac{k_N^2 + 2k_F^2}{2}$ (see condition 16 in the main text) imply that the participation constraints $U_F(\overline{\theta}) > 0$ and $\mathcal{U}_N(\overline{\theta}) > 0$ are indeed satisfied by the solution. We have to verify *ex post* that the neglected constraints, namely condition $\widehat{\gamma}^*(\underline{\theta}) < 1$ and $U_N(\underline{\theta}) < S_N(\underline{\theta}) \Leftrightarrow \pi_N(\underline{\theta}) > 0$, are indeed satisfied. The above system is hard to be solved analytically, because it encompasses a third degree polynomial in $x_N(\underline{\theta})$; 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 $\overline{\theta} = \frac{3}{2}$ and $\underline{\theta} = 1$. Then condition (19) in the main text is satisfied and the solution is such that, for firm N, $x_N^*(\underline{\theta}) = 1.089 > x_N^{FB}(\underline{\theta}) = 1$ and $x_N^*(\overline{\theta}) = x_N^{FB}(\overline{\theta}) = \frac{2}{3}$. Moreover, $U_N^*(\overline{\theta}) = 0.017094$ and $U_N^*(\underline{\theta}) = 0.313\,57$. For firm F, instead, $x_F^*(\underline{\theta}) = x_F^{FB}(\underline{\theta}) = 2$ and $x_F^*(\overline{\theta}) = x_F^{FB}(\overline{\theta}) = \frac{4}{3}$, with $U_F^*(\overline{\theta}) = 0.67521$ and $U_F^*(\underline{\theta}) = 1.1568$. Then, the indifferent worker with high ability has motivation $\widehat{\gamma}^*(\underline{\theta}) = U_F(\underline{\theta}) - U_N(\underline{\theta}) = 1.1568 - 0.31357 = 0.843\,23$, which is higher than that of the indifferent worker with low-ability $\widehat{\gamma}^*(\overline{\theta}) = U_F(\overline{\theta}) - U_N(\overline{\theta}) = 0.67521 - 1.7094 \times 10^{-2} = 0.658\,12$, in line with negative selection of ability for firm N. Finally, wages paid by firm N are $w_N^*(\underline{\theta}) = 0.906\,53$ and $w_N^*(\overline{\theta}) = 0.350\,43$ whereas wages paid by firm F are given by $w_F^*(\underline{\theta}) = 3.1568$ and $w_F^*(\overline{\theta}) = 2.0085$ with $w_i^*(\underline{\theta}) > w_i^*(\overline{\theta})$ for each $i = N, F$. For the sake of comparison, the benchmark contracts in this case would be characterized by $U_N^B(\underline{\theta}) = \frac{1}{3} > U_N^*(\underline{\theta})$ and $U_N^B(\overline{\theta}) = 0 < U_N^*(\overline{\theta})$ for firm N and by $U_F^B(\underline{\theta}) = \frac{7}{6} = 1.166\,7 > U_F^*(\underline{\theta})$ and $U_F^B(\overline{\theta}) = \frac{2}{3} < U_F^*(\overline{\theta})$ for firm F, whereby $\widehat{\gamma}^B(\underline{\theta}) = \frac{5}{6} = 0.833\,33 < \widehat{\gamma}^*(\underline{\theta})$ and $\widehat{\gamma}^B(\overline{\theta}) = \frac{2}{3} > \widehat{\gamma}^*(\overline{\theta})$. Thus, with respect to the benchmark case, for firm N, the labor supply of low-ability workers goes down while the labor supply of high-ability workers goes up. As for wages, we have $w_N^B(\underline{\theta}) = \frac{5}{6} = 0.833\,33 < w_N^*(\underline{\theta})$ and $w_N^B(\overline{\theta}) = \frac{1}{3} < w_N^*(\overline{\theta})$, whereas $w_F^B(\underline{\theta}) = \frac{19}{6} = 3.166\,7 > w_F^*(\underline{\theta})$ and $w_F^B(\overline{\theta}) = 2 < w_F^*(\overline{\theta})$, so that all wages increase at the incentive contracts except for high-ability workers employed by the for-profit firm. Finally, $w_F^*(\underline{\theta}) - w_N^*(\underline{\theta}) = 3.1568 - 0.906\,53 = 2.250\,3 < w_F^B(\underline{\theta}) - w_N^B(\underline{\theta}) = \frac{19}{6} - \frac{5}{6} = \frac{7}{3} = 2.333\,3$ and $w_F^*(\overline{\theta}) - w_N^*(\overline{\theta}) = 2.0085 - 0.350\,43 = 1.658\,1 < w_F^B(\overline{\theta}) - w_N^B(\overline{\theta}) = 2 - \frac{1}{3} = \frac{5}{3} = 1.666\,7$. So the non-profit wage penalty decreases for all types of workers with respect to the benchmark contracts.

2 Negative selection of ability into firm N: Optimal contracts when UIC_N and DIC_F bind

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

Consider now problem (P_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

$$\frac{\partial \mathcal{L}_F}{\partial x_F(\underline{\theta})} = \nu (k_F - \underline{\theta} x_F(\underline{\theta}))(U_F(\underline{\theta}) - U_N(\underline{\theta})) = 0 \tag{15}$$

$$\frac{\partial \mathcal{L}_F}{\partial x_F(\overline{\theta})} = (1-\nu)(k_F - \overline{\theta} x_F(\overline{\theta}))(U_F(\overline{\theta}) - U_N(\overline{\theta})) - \lambda_F^D (\overline{\theta} - \underline{\theta}) x_F(\overline{\theta}) = 0 \tag{16}$$

$$\frac{\partial \mathcal{L}_F}{\partial U_F(\underline{\theta})} = \nu \left(k_F x_F(\underline{\theta}) - \tfrac{1}{2}\underline{\theta} x_F^2(\underline{\theta}) - U_F(\underline{\theta})\right) - \nu (U_F(\underline{\theta}) - U_N(\underline{\theta})) + \lambda_F^D = 0 \tag{17}$$

$$\frac{\partial \mathcal{L}_F}{\partial U_F(\overline{\theta})} = (1-\nu)\left(k_F x_F(\overline{\theta}) - \tfrac{1}{2}\overline{\theta} x_F^2(\overline{\theta}) - U_F(\overline{\theta})\right) - (1-\nu)(U_F(\overline{\theta}) - U_N(\overline{\theta})) - \lambda_F^D = 0 \tag{18}$$

From (15) and (16) 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

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

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 1, whereby $x_N^*(\underline{\theta}) > x_N^{FB}(\underline{\theta})$ and $x_N^*(\overline{\theta}) = x_N^{FB}(\overline{\theta})$. Thus, the following chain of inequalities holds with respect to optimal effort levels

$$x_F^*(\underline{\theta}) = x_F^{FB}(\underline{\theta}) > x_F^*(\overline{\theta}) > x_N^*(\underline{\theta}) > x_N^*(\overline{\theta}) = x_N^{FB}(\overline{\theta}). \tag{19}$$

As for utilities, from (17), 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. Indeed, consider condition (17), substitute $U_N(\underline{\theta})$ for (13) and solve 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 benchmark 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

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

a necessary condition being that $2\lambda_F^D > \lambda_N^U$. Similarly, take condition (18) and, using (12), solve it for $U_F(\overline{\theta})$, yielding

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

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

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

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 leaves an information rent to high-ability workers to prevent them from mimicking low-ability types; this fact also implies that $2\lambda_F^D > \lambda_N^U$ and that $U_F^*(\overline{\theta}) < U_F^B(\overline{\theta})$ must also holds true.

Analyzing now the selection effects, take the analogue of condition (17) at the benchmark, i.e. with $\lambda_F^D = 0$, substitute for $x_F^{FB}(\underline{\theta})$ and solve for the first-best 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 $S_F^{FB}(\underline{\theta})$ into (17), and taking into account that $\lambda_F^D > 0$ in this case, one obtains

$$2\left(U_F^*(\underline{\theta}) - U_F^B(\underline{\theta})\right) - \left(U_N^*(\underline{\theta}) - U_N^B(\underline{\theta})\right) > 0. \tag{20}$$

Considering condition (18) and repeating the same procedure, with the difference that $S_F\left(\overline{\theta}\right) < S_F^{FB}\left(\overline{\theta}\right)$, one gets

$$U_N^*\left(\overline{\theta}\right) - U_N^B\left(\overline{\theta}\right) - 2\left(U_F^*\left(\overline{\theta}\right) - U_F^B\left(\overline{\theta}\right)\right) > 0 \tag{21}$$

Moreover, considering firm N's program and applying the same reasoning to first-order conditions (9) and (10) yields

$$U_F^*(\underline{\theta}) - U_F^B(\underline{\theta}) - 2\left(U_N^*(\underline{\theta}) - U_N^B(\underline{\theta})\right) > 0 \tag{22}$$

and

$$2\left(U_N^*\left(\overline{\theta}\right) - U_N^B\left(\overline{\theta}\right)\right) - \left(U_F^*\left(\overline{\theta}\right) - U_F^B\left(\overline{\theta}\right)\right) > 0, \tag{23}$$

respectively. Finally, putting (20) and (22) together, and rearranging, yields

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

and similarly, putting (21) and (23) 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 and optimal incentive contracts are in place.

Finally, the complete system of equations characterizing the simultaneous solution to both firm's programs 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(\underline{\theta})\right)\right)\left(\underline{\theta} x_N(\underline{\theta}) - 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(\underline{\theta}) & = 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(\underline{\theta})\right)\right) + & \\
+\nu\left(k_N x_N(\underline{\theta}) - \frac{1}{2}\underline{\theta} x_N^2(\underline{\theta}) - 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(\underline{\theta}) - 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(\underline{\theta}) - U_N\left(\overline{\theta}\right) - \frac{k_F^2}{2\underline{\theta}}\right) & = 0
\end{cases}$$

where the relevant unknowns are $x_F(\overline{\theta})$ and $x_N(\underline{\theta})$ together with $U_F(\overline{\theta})$ and $U_N(\overline{\theta})$. Again, one has to verify *ex post* that all neglected constraints, namely the participation constraints $U_F(\overline{\theta}) > 0$ and $\mathcal{U}_N(\overline{\theta}) > 0$, condition $\widehat{\gamma}^*(\underline{\theta}) < 1$ and $\pi_N(\underline{\theta}) > 0$ be indeed satisfied.

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 $\overline{\theta} = \frac{6}{5}$ and $\underline{\theta} = 1$. Then condition (20) in the main text 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 into firm N. Finally, wages paid by firm N are $w_N^*(\underline{\theta}) = 0.90582$ and $w_N^*(\overline{\theta}) = 0.60544$ whereas wages 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 $i = N, F$.

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 of high-ability workers decreases while the labor supply of low-ability workers increases at the incentive contracts. 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 at the incentive contracts, 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.

3 Positive selection of ability into firm N: Optimal contracts when UIC_F binds

Assume that $k_F < k_N$. Consider firm F and assume that UIC_F is binding while DIC_F is slack. The program is (P_F) subject to UIC_F 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

$$\frac{\partial \mathcal{L}_F}{\partial x_F(\underline{\theta})} = \nu (k_F - \underline{\theta} x_F(\underline{\theta}))(U_F(\underline{\theta}) - U_N(\underline{\theta})) + \lambda_F^U (\overline{\theta} - \underline{\theta}) x_F(\underline{\theta}) = 0 \quad (24)$$

$$\frac{\partial \mathcal{L}_F}{\partial x_F(\overline{\theta})} = (1-\nu)\left(k_F - \overline{\theta} x_F(\overline{\theta})\right)\left(U_F(\overline{\theta}) - U_N(\overline{\theta})\right) = 0 \quad (25)$$

where, from (25), 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 (24), 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

$$\frac{\partial \mathcal{L}_F}{\partial U_F(\underline{\theta})} = \nu \left(k_F x_F(\underline{\theta}) - \tfrac{1}{2}\underline{\theta} x_F^2(\underline{\theta}) - U_F(\underline{\theta})\right) - \nu (U_F(\underline{\theta}) - U_N(\underline{\theta})) - \lambda_F^U = 0 \quad (26)$$

$$\frac{\partial \mathcal{L}_F}{\partial U_F(\overline{\theta})} = (1-\nu)\left(k_F x_F(\overline{\theta}) - \tfrac{1}{2}\overline{\theta} x_F^2(\overline{\theta}) - U_F(\overline{\theta})\right) - (1-\nu)\left(U_F(\overline{\theta}) - U_N(\overline{\theta})\right) + \lambda_F^U = 0 \quad (27)$$

Substituting $x_F^{FB}(\overline{\theta})$ into (27) yields

$$\lambda_F^U = (1-\nu)\left(U_F(\overline{\theta}) - U_N(\overline{\theta}) - \left(\frac{k_F^2}{2\overline{\theta}} - U_F(\overline{\theta}) \right) \right), \quad (28)$$

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

$$U_F(\overline{\theta}) > \frac{1}{2}\left(\frac{k_F^2}{2\overline{\theta}} + U_N(\overline{\theta}) \right). \quad (29)$$

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

$$\frac{\partial E(\pi_N)}{\partial x_N(\underline{\theta})} = \nu\left(k_N - \underline{\theta} x_N\left(\underline{\theta}\right)\right)\left(1 - \left(U_F\left(\underline{\theta}\right) - U_N\left(\underline{\theta}\right)\right)\right) = 0 \tag{30}$$

$$\frac{\partial E(\pi_N)}{\partial x_N(\overline{\theta})} = (1-\nu)\left(k_N - \overline{\theta} x_N\left(\overline{\theta}\right)\right)\left(1 - \left(U_F\left(\overline{\theta}\right) - U_N\left(\overline{\theta}\right)\right)\right) = 0 \tag{31}$$

$$\frac{\partial E(\pi_N)}{\partial U_N(\underline{\theta})} = \nu\left(k_N x_N\left(\underline{\theta}\right) - \frac{1}{2}\underline{\theta} x_N^2\left(\underline{\theta}\right) - U_N\left(\underline{\theta}\right)\right) - \nu\left(1 - \left(U_F\left(\underline{\theta}\right) - U_N\left(\underline{\theta}\right)\right)\right) = 0 \tag{32}$$

$$\frac{\partial E(\pi_N)}{\partial U_N(\overline{\theta})} = (1-\nu)\left(k_N x_N\left(\overline{\theta}\right) - \frac{1}{2}\overline{\theta} x_N^2\left(\overline{\theta}\right) - U_N\left(\overline{\theta}\right)\right) - (1-\nu)\left(1 - \left(U_F\left(\overline{\theta}\right) - U_N\left(\overline{\theta}\right)\right)\right) = 0 \tag{33}$$

Conditions (30) and (31) yield first-best effort levels, whereby $x_F^*(\theta) = \frac{k_F}{\theta} = x_F^{FB}(\theta)$ 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

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

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

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

Using (34), one has

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

so that the following chain of inequalities holds

$$x_F^*\left(\overline{\theta}\right) = x_F^{FB}\left(\overline{\theta}\right) < x_F^{FB}\left(\underline{\theta}\right) < x_F^*\left(\underline{\theta}\right) < \frac{k_N}{\sqrt{\underline{\theta}\overline{\theta}}} < \frac{k_N}{\underline{\theta}} = x_N^*\left(\underline{\theta}\right) = x_N^{FB}\left(\underline{\theta}\right). \tag{35}$$

Notice that $x_N^*\left(\overline{\theta}\right) = x_N^{FB}\left(\overline{\theta}\right)$ is missing from the above chain because its position cannot be determined unambiguously. The effort level $x_N^{FB}\left(\overline{\theta}\right) = \frac{k_N}{\overline{\theta}}$ is lower than $\frac{k_N}{\sqrt{\underline{\theta}\overline{\theta}}}$ and higher than $x_F^{FB}\left(\overline{\theta}\right)$. 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 Section 1 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 (34) and (28) into equations (24) and (26), and considering UIC_F binding, yields a system of two equations in two unknowns $x_F(\underline{\theta})$ and $U_F(\overline{\theta})$, which is the following

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

and its solution defines the optimal screening contracts in this case. The finding $U_i^*(\overline{\theta}) > U_i^B(\overline{\theta})$, together with inequality $\overline{\theta} < \frac{k_N^2+2k_F^2}{2}$ (see condition 17 in the main text) imply that the participation constraints $U_F(\overline{\theta}) > 0$ and $\mathcal{U}_N(\overline{\theta}) > 0$ are indeed satisfied by the solution. We have to verify *ex post* that the neglected constraints, namely condition $\widehat{\gamma}^*(\underline{\theta}) > 0$ and $U_F(\underline{\theta}) < S_F(\underline{\theta}) \Leftrightarrow \pi_F(\underline{\theta}) > 0$, are also satisfied.

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) in the main text is satisfied and the solution is such that, for firm N, $x_N(\underline{\theta}) = x_N^{FB}(\underline{\theta}) = \sqrt{2} = 1.4142$ and $x_N(\overline{\theta}) = x_N^{FB}(\overline{\theta}) = \frac{4\sqrt{2}}{5} = 1.1314$. Moreover $U_N^*(\overline{\theta}) = 0.001615$ and $U_N^*(\underline{\theta}) = 0.16505$. For firm F, instead, $x_F^*(\underline{\theta}) = 1.0074 > x_F^{FB}(\underline{\theta}) = 1$ and $x_F(\overline{\theta}) = x_F^{FB}(\overline{\theta}) = \frac{4}{5} = 0.8$ with $x_N^{FB}(\overline{\theta}) > x_F^*(\underline{\theta})$. Moreover, $U_F^*(\overline{\theta}) = 0.20323$ and $U_F^*(\underline{\theta}) = 0.33009$. Then, the motivation of the high-ability worker who is indifferent between firms is $\widehat{\gamma}^*(\underline{\theta}) = U_F^*(\underline{\theta}) - U_N^*(\underline{\theta}) = 0.33009 - 0.16505 = 0.16504$ which is lower than the motivation of the marginal worker with low-ability which is $\widehat{\gamma}^*(\overline{\theta}) = U_F^*(\overline{\theta}) - U_N^*(\overline{\theta}) = 0.20323 - 0.001615 = 0.20162$ in line with positive selection of ability for firm N. Finally wages paid by firm N firm are $w_N^*(\underline{\theta}) = 1.1651$ and $w_N^*(\overline{\theta}) = 0.80162$ whereas wages paid by firm F are given by $w_F^*(\underline{\theta}) = 0.83752$ and $w_F^*(\overline{\theta}) = 0.60323$ with $w_i(\underline{\theta}) > w_i(\overline{\theta})$ for each $i = N, F$ and $w_N(\theta) > w_F(\theta)$ for each $\theta \in \left\{\underline{\theta}, \overline{\theta}\right\}$. Then non-profit employees experience a wage premium for all ability levels. 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 $U_F(\theta)$ higher than $U_N(\theta)$ 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 $U_N^B(\underline{\theta}) = \frac{1}{6} = 0.16667 > U_N^*(\underline{\theta})$ and $U_N^B(\overline{\theta}) = 0 < U_N^*(\overline{\theta})$ 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 of low-ability workers goes up while the labor supply of 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.

4 Positive selection of ability into firm N: Optimal contracts when UIC_F and DIC_N bind

For firm F, UIC_F is binding while DIC_F is slack. The program (P_F), the Lagrangian associated with it and the first-order conditions are the same as in the preceding case.

Consider firm N and its problem (P_N) subject to the constraint that DIC_N binds. Then, the Lagrangian associated with problem (P_N) is

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

The first-order conditions with respect to effort levels are

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

$$\frac{\partial \mathcal{L}_N}{\partial x_N(\overline{\theta})} = (1-\nu)\left(k_N - \overline{\theta} x_N(\overline{\theta})\right)\left(1 - \left(U_F(\overline{\theta}) - U_N(\overline{\theta})\right)\right) - \lambda_N^D(\overline{\theta} - \underline{\theta})\, x_N(\overline{\theta}) = 0 \tag{37}$$

From (36), 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 (37) one has that

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

In particular,

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

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

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

For firm F, the optimal allocation is such that $x_N^*\left(\underline{\theta}\right)=x_N^{FB}\left(\underline{\theta}\right)$ 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

$$x_N^*\left(\underline{\theta}\right)=x_N^{FB}\left(\underline{\theta}\right)>x_N^{FB}\left(\overline{\theta}\right)>x_N^*\left(\overline{\theta}\right)>x_F^*\left(\underline{\theta}\right)>x_F^{FB}\left(\underline{\theta}\right)>x_F^*\left(\overline{\theta}\right)=x_N^{FB}\left(\overline{\theta}\right). \tag{38}$$

The first-order conditions with respect to utilities are

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

$$\frac{\partial\mathcal{L}_N}{\partial U_N\left(\overline{\theta}\right)}=(1-\nu)\left(k_Nx_N\left(\overline{\theta}\right)-\tfrac{1}{2}\overline{\theta}x_N^2\left(\overline{\theta}\right)-U_N\left(\overline{\theta}\right)\right)-(1-\nu)\left(1-\left(U_F\left(\overline{\theta}\right)-U_N\left(\overline{\theta}\right)\right)\right)-\lambda_N^D=0 \tag{40}$$

Analyzing utilities and following the same logic as in Section 2 it is possible to show that $U_N^*\left(\underline{\theta}\right)>U_N^B\left(\underline{\theta}\right)$ whereas $U_F^*\left(\underline{\theta}\right)<U_F^B\left(\underline{\theta}\right)$ and that $U_F^*\left(\overline{\theta}\right)>U_F^B\left(\overline{\theta}\right)$ whereas $U_N^*\left(\overline{\theta}\right)<U_N^B\left(\overline{\theta}\right)$. Thus, it also happens that

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

and that

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

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 in order to define optimal contracts is given by (37) and (40) for firm N together with (24) and (26) for firm F. Using UIC_F and DIC_N binding, allows us to eliminate $U_F\left(\overline{\theta}\right)$ and $U_N\left(\overline{\theta}\right)$, respectively, from the system thus yielding

$$\begin{cases}
\nu\frac{k_N^2}{2\underline{\theta}}-\left(1-U_F\left(\underline{\theta}\right)+2U_N\left(\underline{\theta}\right)\right)-(1-\nu)\left(\frac{1}{2}\underline{\theta}x_N^2\left(\overline{\theta}\right)-\frac{1}{2}\left(\overline{\theta}-\underline{\theta}\right)\left(x_N^2\left(\overline{\theta}\right)-x_F^2\left(\underline{\theta}\right)\right)-k_Nx_N\left(\overline{\theta}\right)\right) & =0\\
(1-\nu)\left(\left(\overline{\theta}-\underline{\theta}\right)x_F^2\left(\underline{\theta}\right)-\frac{1}{2}\left(\overline{\theta}-\underline{\theta}\right)x_N^2\left(\overline{\theta}\right)+\frac{k_F^2}{2\overline{\theta}}\right)-\left(2U_F\left(\underline{\theta}\right)-U_N\left(\underline{\theta}\right)\right)+\nu\left(k_Fx_F\left(\underline{\theta}\right)-\frac{1}{2}\underline{\theta}x_F^2\left(\underline{\theta}\right)\right) & =0\\
\left(\nu\left(k_F-\underline{\theta}x_F\left(\underline{\theta}\right)\right)+\left(\overline{\theta}-\underline{\theta}\right)x_F\left(\underline{\theta}\right)(1-\nu)\right)\left(U_F\left(\underline{\theta}\right)-U_N\left(\underline{\theta}\right)\right)+ & \\
+\left(\overline{\theta}-\underline{\theta}\right)x_F\left(\underline{\theta}\right)(1-\nu)\left(U_F\left(\underline{\theta}\right)-\left(\overline{\theta}-\underline{\theta}\right)x_F^2\left(\underline{\theta}\right)+\frac{1}{2}\left(\overline{\theta}-\underline{\theta}\right)x_N^2\left(\overline{\theta}\right)-\frac{k_F^2}{2\overline{\theta}}\right) & =0\\
(1-\nu)\left(k_N-\overline{\theta}x_N\left(\overline{\theta}\right)\right)\left(1-U_F\left(\underline{\theta}\right)+U_N\left(\underline{\theta}\right)-\frac{1}{2}\left(\overline{\theta}-\underline{\theta}\right)\left(x_N^2\left(\overline{\theta}\right)-x_F^2\left(\underline{\theta}\right)\right)\right)+ & \\
-\nu\left(\overline{\theta}-\underline{\theta}\right)x_N\left(\overline{\theta}\right)\left(1-U_F\left(\underline{\theta}\right)+2U_N\left(\underline{\theta}\right)-\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 $U_N(\underline{\theta})$. Again, one has to verify *ex post* that all neglected constraints, namely the participation constraints $U_F(\overline{\theta}) > 0$ and $\mathcal{U}_N(\overline{\theta}) > 0$, condition $\widehat{\gamma}^*(\underline{\theta}) > 0$ and $\pi_F(\underline{\theta}) > 0$ are indeed satisfied.

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 $\widehat{\gamma}^*(\overline{\theta}) = U_F^*(\overline{\theta}) - U_N^*(\overline{\theta}) = 0.22667 - 0.027879 = 0.19879$, in line with positive selection of ability for firm N. Finally wages paid by firm N firm are $w_N^*(\underline{\theta}) = 1.1665$ and $w_N^*(\overline{\theta}) = 0.85978$ whereas wages paid by firm F are given by $w_F^*(\underline{\theta}) = 0.83982$ and $w_F^*(\overline{\theta}) = 0.64334$ with $w_i^*(\underline{\theta}) > w_i^*(\overline{\theta})$ for each $i = N, F$ and $w_N^*(\theta) > w_F^*(\theta)$ for each $\theta \in \{\underline{\theta}, \overline{\theta}\}$. Then, as in the previous case, non-profit employees experience a wage premium for all ability levels. 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 $U_F(\theta)$ higher than $U_N(\theta)$ for all $\theta \in \{\underline{\theta}, \overline{\theta}\}$. For the sake of comparison, the benchmark contracts in this case would be characterized by $U_N^B(\underline{\theta}) = \frac{1}{6} = 0.16667 > U_N^*(\underline{\theta})$ and $U_N^B(\overline{\theta}) = \frac{1}{36} = 0.027778 < U_N^*(\overline{\theta})$ for firm N and by $U_F^B(\underline{\theta}) = \frac{1}{3} > U_F^*(\underline{\theta})$ and $U_F^B(\overline{\theta}) = \frac{2}{9} = 0.22222 < U_F^*(\overline{\theta})$ for firm F, whereby $\widehat{\gamma}^B(\underline{\theta}) = U_F^B(\underline{\theta}) - U_N^B(\underline{\theta}) = \frac{1}{3} - \frac{1}{6} = 0.16667 > \widehat{\gamma}^*(\underline{\theta})$ and $\widehat{\gamma}^B(\overline{\theta}) = U_F^B(\overline{\theta}) - U_N^B(\overline{\theta}) = \frac{2}{9} - \frac{1}{36} = \frac{7}{36} = 0.19444 < \widehat{\gamma}^*(\overline{\theta})$. Thus, with respect to the benchmark case, for firm N, the labor supply of high-ability workers goes up while the labor supply of low-ability workers goes down. As for wages, we have $w_N^B(\underline{\theta}) = \frac{7}{6} = 1.1667 > w_N^*(\underline{\theta})$ and $w_N^B(\overline{\theta}) = \frac{31}{36} = 0.86111 > w_N^*(\overline{\theta})$, whereas $w_F^B(\underline{\theta}) = \frac{5}{6} = 0.83333 < w_F^*(\underline{\theta})$ and $w_F^B(\overline{\theta}) = \frac{23}{36} = 0.63889 < w_F^*(\overline{\theta})$, so that, with respect to the benchmark, wages increase for firm F while they decrease for firm N. Finally, $w_N^*(\underline{\theta}) - w_F^*(\underline{\theta}) = 1.1665 - 0.83982 = 0.32668 < w_N^B(\underline{\theta}) - w_F^B(\underline{\theta}) = \frac{7}{6} - \frac{5}{6} = \frac{1}{3} = 0.33333$ and $w_N^*(\overline{\theta}) - w_F^*(\overline{\theta}) = 0.85978 - 0.64334 = 0.21644 < w_N^B(\overline{\theta}) - w_F^B(\overline{\theta}) = \frac{31}{36} - \frac{23}{36} = \frac{2}{9} = 0.22222$. So the non-profit wage premium decreases for all types of workers with respect to the benchmark contracts.