Online Appendix

In this supplementary material, we discuss a few extensions of the theoretical framework presented in the paper.

1. Variability in performance given the effort

Our theoretical results would not be qualitatively different if we consider a model with variability in production given effort. To see this, let us consider the following extension of our basic model in which worker i's production when exerting an effort level of e_i can be $y_i = e_i$ with probability $(1-p)$ or zero with probability $p > 0$.[1] Therefore, his expected production is $(1-p)y_i$ and his expected payoff function is given by:

$$\pi_i^w(A, \mathrm{e}, e_i) = A \cdot (1-p) \cdot M(\mathrm{y}) - c(e_i; \theta_i)$$

Worker i's first order condition is given by:

$$c_e(e_i; \theta_i) = \begin{cases} 0 & if \quad y_i \neq M(\mathrm{y}) \\ A \cdot (1-p) & if \quad y_i = M(\mathrm{y}) \end{cases}$$

As a result, the upper bound in the distribution of possible NE described in Proposition 1 is lower when $p > 0$, but all the results described in the paper still apply.

The intuition why variability in production does not play a role in our results has to do with the fact that moral hazard is not an issue in our environment. The main idea of moral hazard problems in teams where compensation depends on the overall joint production (Holmstrom 1982) is that members of a team can cover improper actions behind the uncertainty concerning who was at fault. However, as emphasized by Van Huyck et al. (1990), the inefficient outcomes of the weakest-link game describe in Proposition 1 are not due to moral hazard. This is because a worker's low effort does not affect anyone's payoffs unless he is the weakest-link in which case it affects him just as much as the rest of the team. Thus, a free rider causes the whole project to collapse and this motivates everyone to choose the same effort regardless of production uncertainty.

2. Alternative specifications of the non-monetary utility function

It might be reasonable to consider that workers care, at least partially, about the team outcome. The reason why, in our main model, we assume that workers' non-monetary utility $v(\cdot)$ is a function of individual performance is that if we instead assume that $v(\cdot)$ is a function of team outcome, workers now face the same strategic uncertainty in the wage function and intrinsic utility function. Therefore, equilibria in that game is similar to the characterization of equilibria given in Proposition 1, when workers only care

[1] Note that this production function can be written as $y_i = \varepsilon e_i$ where ε is a random variable that takes the value 0 with probability p and 1 with probability (1-p). Similar results would be obtained with an additive production function and where continuous random shock, as long as the shock is independent of the worker's effort.

about monetary payoffs. Technically, if we assume that the non-monetary utility is $v(M(\mathrm{y}) - g)$, the only difference will be the upper bound of the set of possible Nash Equilibria -$\tilde{y}$, which in Proposition 1 we proved to be given by $c_y(\tilde{y}; \theta_1) = A$. If workers care about the team outcome instead, the upper bound $\tilde{y}_M$ will be given by $c_y(\tilde{y}_M; \theta_1) = A + v'(\tilde{y}_M - g)$; and hence, $\tilde{y}_M > \tilde{y}$. Thus, when workers derive non-monetary utility from the team outcome the result is equivalent to the case in which only monetary incentives are presents in the sense that all workers exert the same level of effort and the upper bound of possible efforts is determined by the maximum performance of the weak-link worker. Since workers now also derive a non-monetary utility from the team outcome, the weak-link worker (whose performance determines team outcome) has now higher incentives to work than with only monetary incentives, so the upper bound will be higher. In other words, just as in the standard case no worker will perform more than the weak-link does because it will be "wasted effort" (both in the monetary and in the non-monetary utility) since team outcome would be unchanged.

Finally, because the manager's outcome is the same as the workers', this would provide the same reference point in $v(\cdot)$, and we again have the same characterization of equilibria as in Proposition 1. One might also think that workers care about manager's individual performance in the task (since managers in our experiment are also working) and perhaps consider manager's performance as an alternative "goal". Even though this is a very reasonable point, workers are not informed about the manager's performance in the experiment and it does not affect team outcome in any way.

Therefore, only when workers derive non-monetary utility from their own effort relative to the goal we have a NE in which some workers choose to worker beyond the performance of the weak-link as characterized in Proposition 2.

3. <u>Specific functional form</u>

We parameterize the intrinsic motivation and cost functions to develop a specific example with closed form solutions of the weak-link game with goal-dependent non-monetary utility. We hope this will help some readers build intuition for the general model in the main text.

In particular, we consider the following widely used prospect theoretic function that satisfies properties $(i) - (v)$:

$$v(y_i - g) = \begin{cases} (y_i - g)^{\frac{1}{2}} & if \quad y_i > g \\ -\lambda(g - y_i)^{\frac{1}{2}} & if \quad y_i \leq g \end{cases}$$

with $\lambda > 1$.

We also consider the increasing and convex cost of effort function: $c(y_i, \theta_i) = \frac{y_i^{3/2}}{\frac{3}{2}\theta_i}$, where $c_y(y_i, \theta_i) = \frac{y_i^{1/2}}{\theta_i}$.

As we discussed in the main text our model considers a situation where worker's payoff function is given by the sum of monetary payoffs (determined by the weak-link performance) and the non-monetary value function (determined by goal attainment) as well as the cost of effort:

$$\pi_i^W(A, \mathbf{y}, y_i, g) = A \cdot M(\mathbf{y}) + v(y_i - g) - c(y_i; \theta_i),$$

where $M(\mathbf{y}) = min\{y_1, \dots, y_n\}$ is the team production determined by the weak-link performance.

The first order condition of a worker's payoff maximization problem are given by:

$$y_i^{\frac{1}{2}} = A \cdot \theta_i \cdot \mathbf{1}_{y_i = M(\mathbf{y})} + \frac{\theta_i}{2}(y_i - g)^{-\frac{1}{2}} \quad if \quad y_i > g \qquad \text{[O1]}$$

$$y_i^{\frac{1}{2}} = A \cdot \theta_i \cdot \mathbf{1}_{y_i = M(\mathbf{y})} + \frac{\theta_i}{2}\lambda(g - y_i)^{-\frac{1}{2}} \quad if \quad y_i \le g \qquad \text{[O2]}$$

Using the some nomenclature we used in the paper we refer to the optimal work performance when the worker is not the weak-link as $\underline{y}(\theta_i, g)$ and the optimal work performance of a weak-link as $\overline{y}(\theta_i, g)$. As we show in the proof of Proposition 2, $\underline{y}(\theta_i, g) < \overline{y}(\theta_i, g)$ and both $\underline{y}(\theta_i, g)$ and $\overline{y}(\theta_i, g)$ are decreasing in θ_i.

Given our functional forms we are able to specify the two group of workers indicated in Proposition 2: low ability workers who perform at the weak-link level and high ability workers who perform above the weak-link in equilibrium.

Let us define the following threshold for θ:

$$\tilde{\theta} = \left\{\theta : \underline{y}(\theta_i, g) = \overline{y}(\theta_1, g)\right\}$$

Note that $\tilde{\theta}$ divides workers into low and high ability workers. When $\theta_i \le \tilde{\theta}$ then $\underline{y}(\theta_i, g) \ge \overline{y}(\theta_1, g)$ and the worker will exert effort equal to $\underline{y}(\theta_i, g)$. When $\theta_i > \tilde{\theta}$ then $\underline{y}(\theta_i, g) < \overline{y}(\theta_1, g)$ and the worker performs the same as the weak-link does, which under the payoff dominance equilibrium is given by $\overline{y}(\theta_1, g)$, which by notational convenience we refer to $\overline{y}_g$

Moreover, from [O1] and [O2] we get that:

$$\underline{y}(\theta_i, g) = \begin{cases} \frac{g + \sqrt{g^2 + \theta_i^2}}{2} & if \quad \underline{y}(\theta_i, g) > g \\ \frac{g - \sqrt{g^2 - \lambda^2 \theta_i^2}}{2} & if \quad \underline{y}(\theta_i, g) \le g \end{cases} \qquad \text{[B3]}$$

Therefore, the threshold $\tilde{\theta}$ is given by:

$$\tilde{\theta} = \begin{cases} 2\sqrt{\overline{y}_g\left(\overline{y}_g - g\right)} & if \quad \underline{y}(\theta_i, g) > g \\ \frac{2}{\lambda}\sqrt{\overline{y}_g\left(\overline{y}_g - g\right)} & if \quad \underline{y}(\theta_i, g) \leq g \end{cases}$$

Note that if $\lambda = 1$ (i.e., workers are no loss averse), the threshold is the same regardless of whether workers attain the goal or not. However, loss averse agents are more willing to work to decrease the pain of performing below the goal so more agents are willing to perform above the weak-link level. Thus, if $\underline{y}(\theta_i, g) \leq g$ then $\frac{d\tilde{\theta}}{d\lambda} > 0$.[2]

From Equation O3 we can also check that,

$$\frac{d\underline{y}(\theta_i, g)}{dg} = \begin{cases} \frac{1}{2} + \frac{g}{2\sqrt{g^2 + \theta_i^2}} & if \quad \underline{y}(\theta_i, g) > g \\ \frac{1}{2} - \frac{g}{2\sqrt{g^2 - \lambda^2\theta_i^2}} & \underline{y}(\theta_i, g) \leq g \end{cases}$$

Thus, in line with Proposition 3.i., when the goal is attainable performance increases with the goal ($\frac{d\underline{y}(\theta_i,g)}{dg} > 0$ if $\underline{y}(\theta_i, g) > g$) but if the goal is unattainable work performance decrease with the goal ($\frac{d\underline{y}(\theta_i,g)}{dg} < 0$ if $\underline{y}(\theta_i, g) \leq g$).

Moreover, our functional forms allows us to compute the optimal goal, the goal that maximizes the weak-link performance. To do this we follow the same steps that we used in the proof of Proposition 3.ii. We start by defining $\hat{g}$ to be the minimum goal that the weak-link worker would fail to attain and by $\hat{y}$ its correspondent performance.

By deriving both sides of Equation B2 (for the weak-link worker, θ_1) with respect to y we get:

$$(\hat{g} - \hat{y}) = \frac{1}{\hat{y}}\left(\lambda\frac{\theta_1}{2}\right)^{2/3}$$

Manipulating Equation B2 at this point we also get that:

$$(\hat{g} - \hat{y}) = \left(\frac{\lambda\theta_1}{2\left(\sqrt{\hat{y}} - A\theta_1\right)}\right)$$

[2] However, as we show in Proposition 3.ii the optimal goal (g^*) is attainable by all workers, therefore, in equilibrium $\tilde{\theta} = 2\sqrt{\overline{y}_{g^*}\left(\overline{y}_{g^*} - g^*\right)}$, which is independent of loss aversion.

Therefore, by solving this system of equations we get:

$$\hat{y} = \left(\frac{A\theta_1}{1-\left(\lambda\frac{\theta_1}{2}\right)^{2/3}}\right)^2 \quad \text{and} \quad \hat{g} = \hat{y} + \frac{1}{\hat{y}}\left(\lambda\frac{\theta_1}{2}\right)^{2/3}$$

Since $\hat{g}$ is the minimum goal that the weak-link would fail to attain, the optimal goal is infinitesimally lower ($g^* = \hat{g} - \varepsilon$ with $\varepsilon \to 0$) and by continuity:

$$g^* = \hat{y} + \frac{1}{\hat{y}}\left(\lambda\frac{\theta_1}{2}\right)^{2/3}$$

See Figures A1 and A2 in the paper's Appendix for illustrations.

Finally, pilling-up and its effect on wasted performance (Proposition 4 in the paper) are easily illustrated here by the fact that $\frac{d(g^*-\hat{y})}{d\lambda} > 0$. Thus, higher loss aversion implies that increasing the goal above the optimal level further decreases performance of the weak-link (and any other worker not attaining the goal). So just as we illustrate in Figure A3 in the paper's appendix, wasted effort would increase when goals are higher than g^*. Moreover, decreasing the goal also increases the dispersion of performance. As we know from Proposition 3.i. decreasing the goal when the goal is attainable lowers work performance. In Figure A.4 we illustrate that this decrease in performance is more intense among lower ability workers whose performance is closer to the goal. Given our specific functional forms we can easily check this result by deriving Equation O3, when the goal is attained, with respect to g and θ_i

$$\frac{d^2\underline{y}(\theta_i, g)}{dgd\theta_i} = -\frac{g\left(g^2 + \theta_i^2\right)^{-\frac{1}{2}}}{4\left(g^2 + \theta_i^2\right)} < 0$$

Thus goals and worker's ability are substitutes when the goal is attainable. Therefore, decreasing the goal below the optimal level (which is attainable for all) decreases the weak-link's performance more than the performance of higher ability workers, increasing wasted-effort among team members.