Teamwork as a Self-Disciplining Device
Online Appendix

Matthias Fahn and Hendrik Hakenes

November 14, 2018

1 Naive and Partially Naive Agents

The agent in our setup is sophisticated and can perfectly anticipate his future self-control problems, hence his future behavior. In this section, we extend our model and also allow for (partially) naive agents in the sense of O'Donoghue and Rabin (2001): An agent's *actual* self control problems in every period are characterized by β. An agent's belief concerning his self-control problems in all future periods, though, is given by $\hat{\beta}$, with $\beta \leq \hat{\beta} \leq 1$. Previously, we had $\hat{\beta} = \beta$. A fully naive agent has $\hat{\beta} = 1$ and believes that he is going to have no self-control problems in the future. A partially naive agent has $\hat{\beta} \in (\beta, 1)$ and is aware of having self-control problems in the future, but underestimates their degree.

To keep the analysis simple, we assume β and $\hat{\beta}$ to remain constant over time and exclude learning. Hence, although an agent's true β is the same in every period, he thinks that the value in future periods is $\hat{\beta}$. This has a direct impact on an agent's perceptions of future individual production. Although he would choose effort $e^{\mathrm{I}} = \beta\delta V/2$ in every period working on his own, he expects to work harder in the future and then choose $\hat{e}^{\mathrm{I}} = \hat{\beta}\delta V/c \geq e^{\mathrm{I}}$. Therefore, it becomes more difficult to enforce team-effort, and the (IC) constraint becomes

$$\left(\beta\delta e^{\mathrm{T}}\frac{V}{2} - \frac{c\,(e^{\mathrm{T}})^2}{2}\right) - \left(\beta\delta e^{\mathrm{I}} V - \frac{c\,(e^{\mathrm{I}})^2}{2}\right)$$
$$+\frac{\beta\delta}{1-\delta}\Big[\big(\delta e^{\mathrm{T}} V - \frac{c\,(e^{\mathrm{T}})^2}{2}\big) - \big(\delta \hat{e}^{\mathrm{I}} V - \frac{c\,(\hat{e}^{\mathrm{I}})^2}{2}\big)\Big] \geq 0. \tag{IC}$$

Whereas the first line is unaffected by an agent's belief concerning his future self-control problems, the second line is reduced – because having to work on individual projects in the future (incorrectly) seems to be less unattractive for partially naive agents. In the extreme case of fully naive agents ($\hat{\beta} = 1$), no team-effort at all can be enforced, for the same reason that made teamwork impossible for agents without self-control problems: Because agents expect to exert first-best effort if working on their own in the future, they perceive a breakdown of the team to be costless. All this implies that teamwork is in principle feasible with partially naive agents; however, a larger degree of naiveté (a higher $\hat{\beta}$) makes cooperation within teams more difficult to sustain.

2 Asymmetries

Our main analysis restricts agents to be identical. In this section, we briefly present one example of asymmetric agents. We show that an agent *without* self-control problems can be part of a productive team if matched with an agent *with* self-control problems – and that such a setting can be mutually beneficial.

Consider a situation with two agents $i = \{1, 2\}$, where $\beta_1 = 1$ and $\beta_2 < 1$. We know from above that two agents without self-control problems cannot form a team. A team of agents 1 and 2 is potentially feasible, though, and helps to relax the self-control problem of the latter. To see this, take agent 1's (IC) constraint,

$$\left(\delta e_1^{\mathrm{T}} \frac{V}{2} - \frac{c\left(e_1^{\mathrm{T}}\right)^2}{2}\right) - \frac{\delta^2 V^2}{2c} + \frac{\delta}{1-\delta}\left[\left(\delta\left(e_1^{\mathrm{T}} + e_2^{\mathrm{T}}\right)\frac{V}{2} - \frac{c\left(e_1^{\mathrm{T}}\right)^2}{2}\right) - \frac{\delta^2 V^2}{2c}\right] \geq 0. \tag{1}$$

There, a solution only exists for $e_2^{\mathrm{T}} > e_1^{\mathrm{T}}$ (for $e_2^{\mathrm{T}} = e_1^{\mathrm{T}}$, the situation is the same as under symmetric matching where agents without self-control problems cannot profitably form a team). Therefore, the seemingly more diligent agent is actually the "lazy" one who only works hard in order to not lose the other one's goodwill.

Several effort-combinations e_1^{T} and e_2^{T} are potentially feasible. As a particular case, suppose we want to enforce $e_1^{\mathrm{T}} = e^{\mathrm{FB}}$. Then, agent 1's (IC) constraint becomes

$$e_2^{\mathrm{T}} \geq \frac{V}{c} \tag{2}$$

Plugging $e_2^{\mathrm{T}} = \frac{V}{c}$ – as well as $e_1^{\mathrm{T}} = e^{\mathrm{FB}}$ – into agent 2's (IC) constraint yields

$$-1 + \delta\left(1 + \beta_2\delta\left(\delta - \beta_2 - \beta_2\delta + \beta_2^2\delta\right)\right) \geq 0. \tag{3}$$

It can be shown that there are combinations of β_2 and δ that satisfy this condition. Given it can be enforced, such an arrangement would be preferred by both agents compared to individual production (this is implied by the (IC) constraints). Agent 1 would naturally prefer such a match to individual production: he contributes first-best effort, whereas agent 2's effort is inefficiently high – hence 1's costs are the same as under individual production but the success probability is higher. He therefore receives an extra rent for serving as a commitment device for agent 2. Agent 2, on the other hand, would rather prefer to be matched with an agent who also has self-control problems, since then the required "markup" on e^{FB} would be lower.

3 Imperfect Public Monitoring

In this section, we sketch the implications of mutual monitoring being imperfect, assuming that each agent can only observe his own effort but generally not the effort level of the other agent. We show that in principle, teamwork also is feasible in this case, only the critical threshold for the discount factor δ is larger. More precisely, we assume that in addition to the resulting output, both agents observe a signal that gives an imperfect notion of exerted effort. With probability p, the signal is informative and real effort is revealed; with probability $1-p$, the signal is uninformative. In the following, we sketch the condition for a perfect public equilibrium (hence, strategies only condition on publicly observable information) with which e^{FB} can be enforced if the relational contract is *only* based on the effort signal. Naturally, the relational contract would work even better if the output realization was also used as an informative signal (then, a low output in combination with an uninformative effort signal might optimally trigger a suspension of cooperation for a number of periods). But this would only effectively improve the outcome if e^{FB} could not be enforced anyway, and a full characterization of a surplus-maximizing equilibrium with imperfect public monitoring (which furthermore might be non-stationary) is beyond the scope of this paper.

Now, a deviation only triggers a punishment if it is actually detected (which happens with probability p). If it is not detected, partners proceed with teamwork. Therefore,

$$U^D = \beta\delta V\left(\frac{e^{\mathrm{T}}}{2}+e^{\mathrm{I}}\right)-\frac{c\,(e^{\mathrm{I}})^2}{2}+\frac{\beta\delta}{1-\delta}\left[p\big(\delta e^{\mathrm{I}}V-\frac{c\,(e^{\mathrm{I}})^2}{2}\big)+(1-p)\big(\delta e^{\mathrm{T}}V-\frac{c\,(e^{\mathrm{T}})^2}{2}\big)\right],$$

and the (IC) constraint for first-best effort yields the condition

$$\delta \geq \delta^{\mathrm{FB}} = \frac{1-\beta\,(1-\beta)}{\left[1-\beta\,(1-p)-\beta^2\left(p\big(2-\beta\big)-1\right)\right]}.$$

Naturally, $\partial\delta^{\mathrm{FB}}/\partial p < 0$. Furthermore, $\delta^{\mathrm{FB}} = (1-\beta\,(1-\beta))\,/\,(1-\beta^2\,(1-\beta))$ for $p = 1$ (the value derived in our benchmark case), whereas $\delta^{\mathrm{FB}} = 1$ for $p = 0$.

4 General Functions

In this section, we show that our results do not depend on a specific functional form of the cost function. More precisely, we assume that total effort costs are $C(e_t)$, with $C' > 0$, $C'' > 0$, $C'' < \infty$, $C(0) = 0$, and $\lim_{e\to 1} C(e) = \infty$. Expected output still amounts to $e_t V$ and is realized in period $t+1$. We keep a linear output function because having a concave output would make it optimal to work on as many small-scale projects as both (because marginal benefits are high for low effort levels). But even if we imposed specific assumptions to get around this problem, having a concave output function would not yield more generality. As long as utilities are concave, any generality can be delivered via the cost component. Now, individual effort is characterized by $\beta\delta V - C'(e^{\mathrm{I}}) = 0$, first-best effort by $\delta V - C'(e^{\mathrm{FB}}) = 0$. The (IC) constraint still amounts to

$$\begin{aligned}&\left(\beta\delta e^{\mathrm{T}}\frac{V}{2}-C(e^{\mathrm{T}})\right)-\big(\beta\delta e^{\mathrm{I}}V-C(e^{\mathrm{I}})\big)\\&+\frac{\beta\delta}{1-\delta}\Big[\big(\delta e^{\mathrm{T}}V-C(e^{\mathrm{T}})\big)-\big(\delta e^{\mathrm{I}}V-C(e^{\mathrm{I}})\big)\Big]\geq 0.\end{aligned}\tag{IC}$$

In the following, we first show that Lemmas 1 and 2, as well as Proposition 1, go through with a general cost function. Then, we derive an additional result where we show that a smaller β might still increase team performance.

First, we can confirm Lemma 1, i. e., that forming a team is not possible if agents do not have inconsistent time preferences.

Lemma B1: *For $\beta = 1$, no positive effort level can be enforced within a team.*

Proof. Note that in this case, $e^{\mathrm{I}} = e^{\mathrm{FB}}$. The constraint becomes

$$\left(\delta e^{\mathrm{T}}\frac{V}{2} - C(e^{\mathrm{T}})\right) - \left(\delta e^{\mathrm{FB}}V - C(e^{\mathrm{FB}})\right)$$
$$+\frac{\delta}{1-\delta}\Big[\big(\delta e^{\mathrm{T}}V - C(e^{\mathrm{T}})\big) - \big(\delta e^{\mathrm{FB}}V - C(e^{\mathrm{FB}})\big)\Big] \geq 0 \tag{4}$$

Now, e^{FB} maximizes $\delta eV - C(e)$, hence $\big(\delta e^{\mathrm{T}}V - C(e^{\mathrm{T}})\big) - \big(\delta e^{\mathrm{FB}}V - C(e^{\mathrm{FB}})\big) \leq 0$ for any e^{T}; furthermore, $\big(\delta e^{\mathrm{T}}\frac{V}{2} - C(e^{\mathrm{T}})\big) - \big(\delta e^{\mathrm{FB}}V - C(e^{\mathrm{FB}})\big) < 0$ for any e^{T}. Therefore, the left hand side of (IC) is strictly negative for any $e^{\mathrm{T}} \geq 0$. ■

Second, we show that Lemma 2 goes through as well.

Lemma B2: *No effort level $e^T \leq e^I$ can be enforced within a team.*

Proof. Assume that $e^{\mathrm{T}} \leq e^{\mathrm{I}}$. Because $e^{\mathrm{I}} \leq e^{\mathrm{FB}}$ and $\delta eV - C(e^{\mathrm{FB}})ce^2/2$ is increasing for effort levels below e^{FB}, the second line of (IC), $\big(\delta e^{\mathrm{T}}V - C(e^{\mathrm{T}})\big) - \big(\delta e^{\mathrm{I}}V - C(e^{\mathrm{I}})\big) \leq 0$ for $e^{\mathrm{T}} \leq e^{\mathrm{I}}$; because e^{I} maximizes $\beta\delta eV - C(e)$ the first line of the (IC) constraint, $\beta\delta e^{\mathrm{T}}V/2 - C(e^{\mathrm{T}}) - \big(\beta\delta e^{\mathrm{I}}V - C(e^{\mathrm{I}})\big)$, is strictly negative. Therefore, the left hand side of (IC) is strictly negative for $e^{\mathrm{T}} \leq e^{\mathrm{I}}$. ■

Finally, we confirm Proposition 1 and show that if δ is sufficiently large, forming a team is feasible for any $\beta < 1$ and first-best effort e^{FB} might eventually be reached.

Proposition B1: *For every $\beta < 1$ and any effort level $e^T \in (e^I, e^{FB}]$, e^T can be enforced within a team if δ is sufficiently close to 1.*

Proof. The second line of (IC), $\big(\delta e^{\mathrm{T}}V - C(e^{\mathrm{T}})\big) - \big(\delta e^{\mathrm{I}}V - C(e^{\mathrm{I}})\big)$, is strictly positive for any $\beta < 1$ and $e^{\mathrm{I}} < e^{\mathrm{T}} \leq e^{\mathrm{FB}}$. Following Lemmas B1 and B2, the first line of the (IC) constraint is negative, however it is bounded for any $\delta \leq 1$. Hence, for $\delta \to 1$, (IC) is satisfied for any e^{T} with $e^{\mathrm{I}} < e^{\mathrm{T}} \leq e^{\mathrm{FB}}$. ■

Finally, we show that a lower β still has two effects: On the one hand, it directly tightens the (DE) because the future becomes less valuable. On the other hand, it relaxes the (DE) constraint by reducing off-path individual effort levels and consequently players' outside options. Starting from $\beta = 1$ and reducing β, the second effect initially dominates. Therefore, more severe self-control problems might actually improve performance of a team.

Proposition B2: *There exists a unique β^{max}, with $0 < \beta^{max} < 1$, that maximizes the scope for cooperation within a team, i. e., β^{max} maximizes the left hand side of the (IC) constraint for any potential team-effort $e^T \leq e^{FB}$.*

Proof. First, note that the left hand side of the (IC) constraint is continuously differentiable with respect to β for arbitrary values of $\hat{e}$. The first derivative of the (IC) constraint with respect to β for any e^{T} is

$$\delta\left(e^{\mathrm{T}}\frac{V}{2} - e^{\mathrm{I}}V\right) - \beta\frac{\delta}{1-\delta}\frac{de^{\mathrm{I}}}{d\beta}\left(\delta V - C'(e^{\mathrm{I}})\right) + \frac{\delta}{1-\delta}\left[\left(\delta e^{\mathrm{T}}V - C(e^{\mathrm{T}})\right) - \left(\delta e^{\mathrm{I}}V - C(e^{\mathrm{I}})\right)\right].$$

Using $C'(e^{\mathrm{I}}) = \beta\delta V$, this expression becomes

$$\delta\left(e^{\mathrm{T}}\frac{V}{2} - e^{\mathrm{I}}V\right) - \beta\frac{\delta}{1-\delta}\frac{de^{\mathrm{I}}}{d\beta}\delta\left(1-\beta\right)V + \frac{\delta}{1-\delta}\left[\left(\delta e^{\mathrm{T}}V - C(e^{\mathrm{T}})\right) - \left(\delta e^{\mathrm{I}}V - C(e^{\mathrm{I}})\right)\right] \tag{5}$$

For all e^{T} with $0 < e^{\mathrm{T}} \leq e^{\mathrm{FB}}$, (5) is negative for $\beta \to 1$ (since $de^{\mathrm{I}}/d\beta = \delta V/C''(e^{\mathrm{I}}) > 0$ and $e^{\mathrm{I}} \to e^{\mathrm{FB}}$), and positive for $\beta \to 0$ (since $de^{\mathrm{I}}/d\beta = \delta V/C''(e^{\mathrm{I}}) > 0$, $C'' < \infty$ and $e^{\mathrm{I}} \to 0$). Because (5) is continuous in β, the left hand side of (IC) must attain at least one global maximum for $\beta \in [0, 1]$. To assess the uniqueness of such a maximum, we compute the second derivative of (IC) with respect to β,

$$-\delta\frac{de^{\mathrm{I}}}{d\beta}V - \frac{\delta}{1-\delta}\frac{de^{\mathrm{I}}}{d\beta}\left(\delta V - C'(e^{\mathrm{I}})\right) + \beta\frac{\delta}{1-\delta}\frac{de^{\mathrm{I}}}{d\beta}\frac{de^{\mathrm{I}}}{d\beta}C''(e^{\mathrm{I}}) - \frac{\delta}{1-\delta}\frac{de^{\mathrm{I}}}{d\beta}\left[\delta V - C'(e^{\mathrm{I}})\right].$$

Using $C'(e^{\mathrm{I}}) = \beta\delta V$ and $\frac{de^{\mathrm{I}}}{d\beta} = \frac{\delta V}{C''(e^{\mathrm{I}})}$, the second derivative of the (IC) constraint with respect to β becomes

$$\frac{de^{\mathrm{I}}}{d\beta}\frac{\delta}{1-\delta}V\left[3\beta\delta - 1 - \delta\right], \tag{6}$$

where $de^{\mathrm{I}}/d\beta > 0$. The term in squared brackets is negative for $\beta < \frac{1+\delta}{3\delta}$ and positive for $\beta > \frac{1+\delta}{3\delta}$, hence – as a function of β – changes its sign exactly once. This allows us to show that the left-hand side of (IC) does not attain a local minimum on $\beta \in (0, 1)$.

For a proof by contradiction, assume there is a minimum at $\beta^{\min} \in (0,1)$. Then,

$$\frac{de^{\mathrm{I}}(\beta^{\min})}{d\beta} \frac{\delta}{1-\delta} V \left[3\beta^{\min}\delta - 1 - \delta\right] \geq 0.$$

If (6) is positive at $\beta^{\min}$, it also must be strictly positive for all $\beta > \beta^{\min}$. Therefore, there cannot be a maximum at any $\beta > \beta^{\min}$, which however would contradict that (5) is negative for $\beta \to 1$. Because the left-hand side of (IC) must attain at least one maximum, there exists a $\beta^{\max}$ where (5) equals zero and where (6) is non-positive. Finally, $\beta^{\max}$ is unique because of the non-existence of a minimum. ■

References

O'Donoghue, T., and M. Rabin (2001): "Choice and Procrastination," *Quarterly Journal of Economics*, 116(1), 121–160.