Fahn, Matthias; Hakenes, Hendrik

Working Paper

Teamwork as a Self-Disciplining Device

Discussion Paper, No. 42

Provided in Cooperation with:
University of Munich (LMU) and Humboldt University Berlin, Collaborative Research Center Transregio 190: Rationality and Competition

Suggested Citation: Fahn, Matthias; Hakenes, Hendrik (2017) : Teamwork as a Self-Disciplining Device, Discussion Paper, No. 42, Ludwig-Maximilians-Universität München und Humboldt-Universität zu Berlin, Collaborative Research Center Transregio 190 - Rationality and Competition, München und Berlin

This Version is available at:
https://hdl.handle.net/10419/185712

Teamwork as a Self-Disciplining Device

Matthias Fahn (LMU Munich and CESifo)
Hendrik Hakenes (University of Bonn and CEPR)

Discussion Paper No. 42

July 27, 2017

Collaborative Research Center Transregio 190 | www.rationality-and-competition.de
Ludwig-Maximilians-Universität München | Humboldt-Universität zu Berlin
Spokesperson: Prof. Dr. Klaus M. Schmidt, University of Munich, 80539 Munich, Germany
+49 (89) 2180 3405 | info@rationality-and-competition.de

Teamwork as a Self-Disciplining Device*

Matthias Fahn† Hendrik Hakenes‡

July 13, 2017

Abstract

We show that team formation can serve as an implicit commitment device to overcome problems of self-control. If individuals have present-biased preferences, effort that is costly today but rewarded at some later point in time is too low from the perspective of an individual's long-run self. If agents interact repeatedly and can monitor each other, a relational contract involving teamwork can help to improve performance. The mutual promise to work harder is credible because the team breaks up after an agent has not kept this promise – which leads to individual underproduction in the future and hence a reduction of future utility.

Keywords: Self-Control Problems, Teamwork, Relational Contracts.

JEL-Classification: L22, L23.

*We thank Reginal Seibel for excellent research assistance. We thank Daniel Barron, Kristina Czura, Florian Englmaier, Daniel Garrett, Bob Gibbons, Marina Halac, Fabian Herweg, Thomas Kittsteiner, Svetlana Katolnik, Daniel Krähmer, Matthias Kräkel, David Laibson, Yves Le Yaouanq, Jin Li, Takeshi Murooka, Muriel Niederle, Martin Peitz, Mike Powell, Ariel Rubinstein and Klaus Schmidt, as well as seminar participants at the University of Hannover, the LMU Munich, the 17th Colloquium on Personnel Economics (Cologne), the 18th SFB/TR 15 meeting (Mannheim), the 18th Annual Conference of The International Society for New Institutional Economics (Duke), the 41st Annual Conference of the European Association for Research in Industrial Economics (Milan), the 2014 Annual Conference of the Verein für Socialpolitik (Hamburg), and the 15th Annual Conference of the German Economic Association of Business Administration (Regensburg) for helpful comments. Matthias Fahn gratefully acknowledges financial support from the German Science Foundation (DFG) through collaborative research center CRC TRR 190. All errors are our own.

†University of Munich and CESifo, Department of Economics, Kaulbachstr. 45, 80539 München, Germany, matthias.fahn@econ.lmu.de.

‡University of Bonn and CEPR. University of Bonn, Institute of Finance and Statistics, Adenauerallee 24-42, 53113 Bonn, hakenes@uni-bonn.de.

Remember teamwork begins by building trust. And the only way to do that is to overcome our need for invulnerability.

Patrick Lencioni, The Five Dysfunctions of a Team: A Leadership Fable

1 Introduction

Teams are formed in all kinds of circumstances. They can be found within firms to address complicated problems (almost all US firms use teams in one form or the other (Lazear and Shaw, 2007)). Lawyers or doctors form partnerships. Potential entrepreneurs start a firm with friends instead of pursuing their ideas alone. Due to its importance, economists have widely analyzed teamwork, mainly focussing on two conflicting aspects. On the one hand, technological benefits and specialization render teamwork necessary in situations that involve complex or risky tasks. On the other hand, teamwork is associated with a free-rider problem. Because the team's common goal is a public good, an underprovision of contributions can result (see Alchian and Demsetz, 1972). Starting with Holmstrom (1982) – who shows in a static setting that the first-best is impossible to reach if a balanced-budget rule is imposed – the literature has tried to identify ways to overcome this public-good problem. More recently, non-technological benefits of teamwork have come into focus. For example, mutual monitoring and peer pressure can foster cooperation within a team and consequently increase productivity (see Kandel and Lazear, 1992 or Baron and Kreps, 1999).

Our paper derives another inherent benefit of teams: Driven by repeated interaction and mutual monitoring, teamwork can help to overcome problems of self-control. If individuals have present-biased preferences, any effort that is costly today but rewarded at some later point in time is too low from the perspective of an individual's long-run self. Then, agents can form a team, share the output generated by their effort, and promise each other to work hard using a relational contract. That way, they can improve their performance – even though the standard free-rider problem is present, and even though teamwork renders no technological benefits. The mutual promise to work hard is credible: after an agent has not kept his promise, the team breaks up, and the agent is again alone with his self-control problems.

More precisely, we propose an infinite-horizon model with two present-biased agents who can repeatedly work on individual projects. Since production is costly today but rewards are realized one period later, an agent works less hard than he would prefer from the perspective of his long-run self. However, forming a team can serve as an endogenous commitment device to increase individual effort levels – provided agents are aware of their own time-inconsistency, hence are sophisticated in the sense of Laibson (1997) and O'Donoghue and Rabin (1999). In a team, agents jointly work on a project, share potential benefits, and make a mutual promise to work hard. Since effort is not verifiable but can only be observed by one's team-mate, the promise to work hard in a team has to be credible. There, an agent faces the following trade-off. On the one hand, he has an incentive to deviate – by slacking off because of his present bias, but also by contributing nothing to the team, free-ride on the other agent's effort, and instead work on an individual project. On the other hand, such a deviation is followed by a reversion to individual production in all subsequent periods. This is costly for the agent because future effort under individual production is regarded too low from an agent's present perspective. Therefore, a deviation entails current gains but future losses. The agent only sticks to the team arrangement if the difference between future utilities obtained in a team and future utilities from individual production is sufficiently large (from today's perspective). This difference depends on the standard discount factor as well as on an agent's present bias. If the standard discount factor is large enough, that is, if the future is rather valuable, the efficient effort level (from the perspective of earlier periods) can be implemented in a team. The effect of an agent's present bias on his incentive to stick to the team agreement is twofold. On the one hand, more severe self-control problems reinforce discounting and consequently increase the incentive to deviate. On the other hand, more severe self-control problems mitigate an agent's incentive to deviate because they reduce effort under individual production and consequently future utilities after a deviation. The latter effect dominates if self-control problems have initially been moderate. In this case, more severe self-control problems of its members actually improve team performance.

Summing up, we show theoretically that working in a team can serve as a self-disciplining device for individuals to overcome their self-control problems. As an illustrative example for the validity of this result, take a scientist's daily work on research projects. Many distractions keep him from being focused and motivated – in particular since most of the rewards of doing research are not realized immediately.

There are ways to increase his commitment, like conference deadlines or tools that temporarily block access to distracting websites. One of the most widespread remedy to tackle motivational issues, though, is the collaboration with co-authors. Besides making use of mutual comparative skills, spurring creativity, and plenty of other advantages, such a cooperation can also serve as a commitment device to overcome self-control problems. Promises made to co-authors are motivating, in particular if one also wants to work with them on future projects.

Concerning more systematic evidence on the role of teamwork as a self-disciplining device, note that our argument relies on two aspects. First, the *reason* for entering a team in order to use it as a commitment device lies in an agent's self-control problems. Second, the *mechanism* to enforce higher effort within a team is a relational contract between its members (in contrast to, for example, peer pressure). This relational contract requires mutual monitoring and repeated interaction (as introduced by Che and Yoo (2001)).

Landeo and Spier (2015) provide experimental evidence for the latter. They show that if agents work in a team, repeated interaction and mutual monitoring ceteris paribus (that is, for a given potential reward) lets them work harder. Bauer, Chytilova, and Morduch (2012) provide evidence from the world of microcredit borrowing that individuals might join teams to overcome problems of self control. They conduct a series of "lab-in-the-field" experiments in South India. Participants have the option to borrow relatively large, real stakes, either from a microcredit institution or other sources. Borrowing from a microcredit institution involves joining a lending group, which implies joint liability (i.e., the whole group is responsible for the loan repayment), weekly group meetings and public transactions. Being part of a lending group hence has many properties of teamwork in our setting – the free-rider problem, mutual monitoring and repeated interaction (repaying a loan immediately gives access to a new, often larger, loan) – whereas borrowing from other sources rather relates to individual production. Bauer, Chytilova, and Morduch (2012) find that among women who borrow, those with present-biased preferences are particularly likely to borrow from microcredit institutions. They argue that these findings support the hypothesis that group-based lending might also serve as a commitment device to overcome problems of self-control.

In a number of extensions we explore the robustness of our results. In the benchmark case, teamwork renders a free-rider problem but no technological benefits.

Therefore, teamwork is not possible with time-consistent agents, because effort when working on an individual project is already at its first-best from the perspective of *any* period. If teamwork is associated with technological benefits (like economies of scale), though – implying that also time-consistent agents would rather work within a team than pursuing individual projects – agents with present-biased preferences might actually perform *better* than agents without. This is again driven by the lower future utilities of present-biased agents under individual production. We also allow for agents to be partially naive with respect to their future self-control problems and to underestimate their magnitude. This makes it more difficult to enforce team-effort because having to work on individual projects in the future is perceived less unattractive from today's perspective. Further extensions can be found in an Online Appendix. There, we first derive a renegotiation-proof equilibrium, with respect to a potential renegotiation of the initial agreement as well as renegotiation of the punishment following a deviation. Second, we show that a mixed team consisting of an agent with self-control problems and an agent without can be used as a commitment device for the former agent. This only works, though, if the agent with self-control problems provides higher effort than the one without; hence, the seemingly lazier agent actually is the one without self-control problems. Third, we sketch the implications of mutual monitoring being imperfect and show that teamwork is also feasible in this case. Fourth, we point out that our results do not depend on a specific functional form of the effort cost function (the results in our paper are derived using a quadratic cost function).

We also discuss the validity of our results. There, we first approach the question whether an agent should be able to play a relational contract not only with another individual, but also with his future selves. Such games have been analyzed by Laibson (1994) or Bernheim, Ray, and Yeltekin (2013), who characterize subgame perfect equilibria of a game played between different selves. We argue that relational contracts in teams not only have a larger intuitive appeal than intra-personal games that require self-inflicted punishment, but also are better suited to prevent renegotiation of the initial agreement. We formally derive the latter argument, slightly extending our model in a sense that the outside option of an agent not only is individual production, but that he also has the option to find a new partner with whom he can form a team. Furthermore, we argue that agents are more likely to actually punish deviations of their partners in teams than their own deviations in intra-personal games. Reasons are that deviations are more salient in teams, and

in particular that individuals often accept excuses for their own behavior too easily. There, we use evidence from psychology stating that a "self-serving bias" or "attribution bias" makes individuals prone to generally accept their own excuses (Zuckerman (1979)). To the contrary, the behavior of others is not excused that easily, which is known as an "actor-observer bias" (Jones and Nisbett (1987, page 80)). Therefore, teamwork can also act as a commitment device to not bring up an excuse if one did not work hard – because one knows that their counterpart will not accept such an excuse and punish them accordingly. Second, we compare our approach to other potential mechanisms to enforce cooperation within a team, peer pressure and reciprocity, and discuss how those could be distinguished empirically. There, we again have to differentiate between the reason for team formation (in our case: an agent's self-control problems) and the mechanism to enforce high effort within a team (in our case: a relational contract). Peer pressure and reciprocity rather represent mechanisms that allow to enforce high effort even in the absence of repeated interaction. Therefore, our approach is likely to be relevant if repeated interaction increases the performance of teams – a prediction that in the lab has already been confirmed by Landeo and Spier (2015). Furthermore, forming a team is always optimal in our case provided it allows to implement high effort. This does not necessarily hold if individuals who suffer from self-control problems contemplate forming a team – but if the mechanism to enforce high effort either is peer pressure or reciprocity.

Related Literature. This paper contributes and relates to three strands of literature; incentives in teams and professional partnerships, present-biased preferences, and relational contracts. The optimal provision of incentives in teams has been widely analyzed (starting with Alchian and Demsetz (1972) and Holmstrom (1982)). This literature, though, mainly assumes that teams are formed exogenously and only joint performance schemes are feasible. More recently, a number of more recent papers have shown that the underlying free-rider problem can be overcome if team members are able to (partially) observe the performance of their peers and hence form relational contracts with each other. Che and Yoo (2001) show that given a team is formed exogenously, joint performance evaluation might be optimal even though the principal observes individual performance signals. The resulting free-rider problem can be overcome by mutual monitoring, repeated interaction and a relational contract formed between agents. Kvaløy and Olsen (2006) extend Che

and Yoo's paper, assuming that the (imperfect) signal the principal receives is not verifiable as well, hence the relationship between principal and agents also is governed by relational contracts. They identify instances for which relative performance evaluation (compared to joint and independent performance evaluation) is optimal and show that this depends on the interaction between agents' discount factors and their productivities. Rayo (2007) derives optimal asset ownership if a verifiable joint performance scheme exists but relational contracts between agents are feasible.[1] Whereas these articles take the existence of teams as exogenously given, the present paper shows that the insights developed there can also be used to derive benefits of endogenous team formation. Furthermore, unlike Che and Yoo (2001) and subsequent papers, we do not analyze relational contracts between a principal several agents, but assume that two identical individuals interact.

Different from the mechanism applied in our paper, the literature has identified further instances where the endogenous formation of teams or partnerships can be optimal. Itoh (1991) shows that teamwork may induce agents to help each other. Bar-Isaac (2007) develops a reputational model where it can be optimal to form a team in order to maintain reputational incentives for older workers who want to sell a firm but whose personal reputation is not at stake anymore. Corts (2007) shows that teamwork can help to overcome multitasking problems, by grouping tasks with a lower and those with a higher impact on observable signals. Mukherjee and Vasconcelos (2011) extend Corts' model by assuming that observable signals are not verifiable. Because teamwork requires higher maximum payments, it is also associated with a higher reneging temptation. Hence, teamwork only works if a firm's discount factor is sufficiently large. More related to our paper, Battaglini, Benabou, and Tirole (2005) illustrate how peer pressure in groups can help individuals to resist temptation and restrain themselves – via learning about one's own self-confidence in relatively homogeneous groups. Extending the literature on endogenous team formation, and using relational contracts between team members as well, we show that teamwork can also enhance productivity if individuals have self-control problems.

Furthermore, we contribute to the literature on inconsistent time preferences and

[1]Several articles derive derive different mechanisms that may render joint performance schemes optimal. Mohnen, Pokorny, and Sliwka (2008) and Bartling (2011) show that if players have social preferences, their preferences for equal outcomes can channel incentives in a way to overcome the free-rider problem. Kim and Vikander (2013) show that if teamwork has *decreasing* returns to scale and relational contracts are used to motivate employees, joint-performance systems can be optimal because they help to smooth payments over time.

self-control problems. Strotz (1955) is the first to formalize this aspect by noting that an individual's discount rate between two periods might depend on the time of evaluation. He further discusses differences between those who recognize this inconsistency – and hence might try to bind their future selves – and those who do not. Phelps and Pollak (1968) state that in particular growth models should take the possibility of inconsistent time preferences into account as this affects savings. Laibson (1997) shows that illiquid assets can serve as a commitment device to bind future selves. O'Donoghue and Rabin (1999) focus on the distinction between individuals who are aware of their time inconsistency and those who are not; they label the former "sophisticated" and the latter "naive".

A large amount of evidence confirms that people make decisions that are not consistent over time, for example when using credit cards or signing up for health clubs (DellaVigna and Malmendier, 2004, 2006). Ashraf, Karlan, and Yin (2006) conduct a field experiment with customers of a Philippine bank, allowing individuals to choose a commitment device that restricts access to their savings. More than 25% of customers opt for this device and subsequently increase their savings substantially. More recently, experimental evidence from the field and the lab uses real-effort tasks to directly identify self-control problems. Kaur, Kremer, and Mullainathan (2010, 2013) perform a field experiment involving full-time workers in an Indian data entry firm. Quantity and quality of output can be easily measured, and workers receive a piece rate. The existence of self-control problems is supported by the observation that workers increase effort as the payday gets closer. In addition, many workers select an offered commitment device that would be dominated for individuals with exponential preferences. Furthermore, Augenblick, Niederle, and Sprenger (2013) perform a real-effort task lab experiment. There, participants show a significant present-bias as well, and many of them demand a binding commitment device if it is offered. We contribute to this literature showing that by forming a team, individuals can create an *implicit* commitment device.[2] Thereby, they use the benefits of future cooperation as a collateral to overcome self-control problems. In addition, we show that people with present-biased preferences can actually perform better than those

[2]Several other implicit commitment devices to overcome self-control problems, in particular to enforce optimal consumption and savings decisions, have been identified. Bond and Sigurdsson (2013) show that contractual arrangements restricting an individual's intertemporal consumption choice can help to solve the trade-off between inducing future commitment and reacting flexibly to stochastic and non-verifiable shocks. Basu (2011) derives a justification for so-called rotational savings and credit associations, which many people in the developing world join. Although clearly restricting an individual's flexibility, they can foster commitment to accumulate savings.

without. To our knowledge, we are the first to derive such a result. It is driven by individuals with self-control problems being hurt more by a breakdown of teamwork.

Finally, we relate to the literature on relational contracts. Relational contracts are implicit arrangements based on observable but non-verifiable information. Theoretical foundations have been laid by Bull (1987) and MacLeod and Malcomson (1989) and later extended for the case with imperfect public monitoring by Levin (2003). This triggered various developments of the baseline model, thereby providing many explanations for real-world phenomena. We show that adding behavioral assumptions to a relational contracting framework can yield new and interesting implications.

2 Basic Model – Individual Production

2.1 The Economy

Consider two risk-neutral agents $i = \{1, 2\}$ who live for infinitely many periods, $t \in \{0, 1, \ldots\}$. Each agent has access to an inexhaustible amount of projects. At each date, an agent chooses a total effort level e_t and how to allocate it among projects (we add an index for a specific agent whenever necessary).

Each project returns V in the following period $(t + 1)$ with probability identical to the effort allocated to this project. The total expected return from effort e_t is thus $e_t V$, independent of the concrete allocation of effort among projects. Hence, an agent can influence his payoff in period $t + 1$ by increasing his effort in the previous period t. Effort leads to an immediate cost $ce_t^2/2$ at date t, with $c > 0$, where e_t is an agent's *total* effort.[3] There are no technological linkages of projects across periods. The effort spent on a project in period t does not affect the likelihood that the project is successful in any later period. If an agent finishes one project, or abandons it, he can start a new project.

[3]Hence, an agent's effort cost and expected payoff is the same, no matter whether he just works on one or allocates total effort among an arbitrary number of different projects. Therefore, the number of projects an agent works on in a given period can be set to 1.

2.2 Discounting

Agents discount future costs and future utilities in a quasi-hyperbolic way according to Laibson (1997) and O'Donoghue and Rabin (1999). Utilities after t periods are discounted with a factor $\beta\delta^t$, with β and δ in $(0;1]$. Hence, the discounted value of a utility stream evaluated in period t is $u_t + \beta\left[\delta u_{t+1} + \delta^2 u_{t+2} + \ldots\right]$, where u_t is the agent's period-t utility. Consequently, an agent's preferences are dynamically inconsistent. At date $t = 0$, an agent would pay $\beta\delta$ for a dollar at date $t = 1$, and at date $t = 1$ he would pay $\beta\delta$ for a dollar at date $t = 2$. However, at date $t = 0$, he would give up $\beta\delta^2$ instead of $\beta^2\delta^2$ for a dollar at date $t = 2$. In addition, we assume that agents are sophisticated in the sense that they are fully aware of their time-inconsistency and hence take their future time-inconsistency into account when taking actions[4]. Throughout the paper, we assume that there is no formal device for an agent to commit to any specific effort level. Furthermore, to make sure that we always have an interior solution, we assume that $\delta V/c < 1/2$.

2.3 Individual Production and Self-Control Problems

Now, we derive effort levels if agents work on their own. There, an agent decides how much he wants to work at the beginning of any period t, maximizing his discounted utility[5]

$$\beta\delta e_t V - \frac{c e_t^2}{2}. \tag{1}$$

The optimal effort level under individual production, e^{I}, is the same in every period and equals

$$e^{\mathrm{I}} = \frac{\beta\delta V}{c}. \tag{2}$$

However, reasoning over how much effort he wants to spend in the future, an agent would come to a different result. Thinking at date t how much he wants to work at

[4]We analyze the case of partially naive agents below, in Section 4.2.

[5]From a formal point of view, one could also consider an agent playing a game with his future selves, conditioning current actions on past behavior. Here, we exclude such intra-personal games and refer to Section 5.1 for a discussion of those.

a future date $\hat{t} > t$, he rather maximizes

$$\beta\delta^{\hat{t}-t}\big(\delta e_{\hat{t}}V - \frac{ce_{\hat{t}}^2}{2}\big). \tag{3}$$

For any period $\hat{t} > t$ this is solved by first-best effort e^{FB}, i. e., by

$$e^{\mathrm{FB}} = \frac{\delta V}{c}. \tag{4}$$

Informally speaking, the agent is lazy. Since $e^{\mathrm{I}} < e^{\mathrm{FB}}$ for $\beta < 1$, he works less than he would originally have liked to work from the perspective of earlier periods. This does not come as a surprise, though, since the agent is sophisticated and fully aware of his self-control problem.

This time-inconsistency problem is not present in period $t = 0$, the first period of the game. There, no past plans do yet exist from which current behavior can deviate. Hence, first-best effort in period $t = 0$ is equal to e^{I}. This phenomenon – that optimal behavior in period $t = 0$ differs from optimal behavior in all subsequent periods – will also be manifest in the description of equilibrium team arrangements.

3 Teamwork

In this section, we discuss exactly the same economic setting as before. The only difference is that – instead of working on their own – agents can form a team at the beginning of every period. This implies that both agents jointly work on some projects, and that the payoff V from these projects – if successful – is shared equally.

Equivalently to footnote 3, the number of projects pursued in a team at each date can – without loss of generality – be set to 1. Agents make their effort choices simultaneously, where effort is mutually observable but not verifiable. Given agent 1 chooses effort $e_{1,t}$ and agent 2 chooses $e_{2,t}$, the joint expected payoff – realized in period $t + 1$ – is $(e_{1,t} + e_{2,t})\,V$, and each expects to receive $(e_{1,t} + e_{2,t})\,V/2$. Hence, we assume that there are no economies (or diseconomies) of scale (or scope) from teamwork[6] – the same amount of work can get done and costs of effort are the same.

[6] In Section 4.1, we introduce economies of scale in a team.

Even after agents have formed a team in a period t, they are always able to revert to individual production, in period t as well as in all future periods. This implies that we rule out exclusivity contracts where agents commit to profit-sharing agreements involving all of an agent's potential projects.

Our definition of teamwork – that agents jointly work on a project – is solely made for concreteness. Any arrangement where one agent uses his effort in order to benefit the other agent would yield the same qualitative results. Hence, our mechanism can also be applied to settings that go beyond a narrow definition of teamwork. For example, one agent might directly spend some of his working time on one of the other agent's projects, and vice versa. Agents could also focus on different topics and explain their insights to each other. Plain profit sharing would also be feasible, as well as any combination of these aspects (like sharing the outputs of two projects and alternate working on it).

3.1 Relational Contracts and Equilibrium Concept

Agents can form a relational contract specifying effort levels within the team. Since both agents can observe each other's effort, mutual monitoring is feasible. This relational contract is formed at the beginning of the game. For any period t, it specifies the actions both agents are supposed to take along the whole path of the game – contingent on the realized history up to period t. The relational contract implicitly determines when a team is supposed to be formed, as well as each agent's effort level on and off the equilibrium path. Both agents' contingent action plans, i. e. their strategies, have to be optimal for any feasible history, i. e., form a subgame perfect equilibrium of the dynamic game. However, given agents' time inconsistency, we require a subgame perfect equilibrium to constitute a Nash equilibrium at each subgame, given agents' preferences once a respective subgame is reached, and given each agent's continuation play.

As explained above, the time-inconsistency problem does not exist in period $t = 0$ (there, first-best effort is identical to e^{I}). Thus, teams can only potentially add value in periods $t \geq 1$. For those periods, relational contracts can be stationary, i. e., team-effort is the same in every period,[7] allowing us to omit time subscripts.

[7] This is because agents are risk-neutral and actions are mutually observed.

3.2 Team Production

Assume each agent is supposed to exert team-effort e^{T} in a period t. For tractability, we focus on symmetric equilibria where effort e^{T} is the same among agents. To explore whether team-effort e^{T} can be enforced, we have to specify what happens after a deviation. In principle, there are two possibilities for an agent to deviate. First, an agent could refuse to form a team-project as specified by the relational contract. Second, after forming the team, the agent could provide an effort level different from e^{T}. The first possibility to deviate, though, is not relevant: Because an agent can always work on his own projects, agreeing to form a team but then choosing $e^{\mathrm{T}} = 0$ is always possible. Therefore, we can restrict our attention to deviations from equilibrium team-effort e^{T}. Given any such deviation, we follow Abreu (1988) who shows that any observable deviation should be responded by the strongest feasible punishment. In our case, that means that cooperation within the relational contract breaks down for good, and agents resume to individual production.

In the following we analyze whether such a relational contract can be sustained as a subgame perfect equilibrium, and in particular whether team effort e^{T} can exceed the effort level of individual production, e^{I} or might even reach e^{FB} – the first-best effort as regarded from the perspective of earlier periods. There, note that e^{FB} also is an agent's preferred symmetric future effort level given a team has been formed: In a period t and contemplating about the preferred effort levels at a future date $\hat{t} > t$ (exerted by both agents), he maximizes

$$\beta\delta^{\hat{t}-t}\left[\delta\left(e_{\hat{t}}^{\mathrm{T}} + e_{\hat{t}}^{\mathrm{T}}\right)V/2 - c\left(e_{\hat{t}}^{\mathrm{T}}\right)^2/2\right], \tag{5}$$

which is solved by e^{FB}. Once a team has been formed and given agents stick to their agreement, an agent's expected discounted utility stream in a period $t \geq 1$ is

$$\begin{aligned} U^{\mathrm{T}} &= \beta\delta e^{\mathrm{T}}V - \frac{c\,(e^{\mathrm{T}})^2}{2} + \sum_{t=1}^{\infty}\beta\delta^t\big(\delta e^{\mathrm{T}}V - \frac{c\,(e^{\mathrm{T}})^2}{2}\big) \\ &= \beta\delta e^{\mathrm{T}}V - \frac{c\,(e^{\mathrm{T}})^2}{2} + \frac{\beta\delta}{1-\delta}\big(\delta e^{\mathrm{T}}V - \frac{c\,(e^{\mathrm{T}})^2}{2}\big). \end{aligned} \tag{6}$$

e^{T} can only be enforced by a relational contract if a deviation is never optimal. Because an agent only gets 0.5 of the outcome of his own effort within the team and because *any* deviation triggers a breakdown of future cooperation, if the agent deviates, he should optimally provide zero team-effort and instead completely work

on an individual project. Thereby he would free-ride on the other agent's team effort. After a deviation, both agents work on individual projects from then on.[8] Hence, an agent's expected discounted utility stream given he joins the team but then deviates is

$$\begin{aligned} U^D &= \beta\delta e^{\mathrm{T}}\frac{V}{2} + \beta\delta e^{\mathrm{I}}V - \frac{c\,(e^{\mathrm{I}})^2}{2} + \sum_{t=1}^{\infty}\beta\delta^t\big(\delta e^{\mathrm{I}}V - \frac{c\,(e^{\mathrm{I}})^2}{2}\big) \\ &= \beta\delta V\left(\frac{e^{\mathrm{T}}}{2} + e^{\mathrm{I}}\right) - \frac{c\,(e^{\mathrm{I}})^2}{2} + \frac{\beta\delta}{1-\delta}\big(\delta e^{\mathrm{I}}V - \frac{c\,(e^{\mathrm{I}})^2}{2}\big). \end{aligned} \tag{7}$$

To sustain teamwork, an agent's on-path utility stream within the team has to be larger than after a deviation. Hence, an incentive compatibility (IC) constraint must be satisfied, $U^{\mathrm{T}} \geq U^D$, or

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

Here, the first line captures the standard free-rider problem of teamwork and is negative; the second line gives the value of future cooperation, evaluated today. Only if the second line dominates, teamwork is feasible. If (IC) is not satisfied, no team is formed, and both agents have utilities U^{I}.

Note that the (IC) constraint must hold in every period t. This implies that – different from many other (formal) commitment devices analyzed in the literature – teamwork has to be optimal for every future self of an agent (taking every future self's continuation utility into account), not only for the period-0 self.

3.3 Results

We now analyze what can be achieved within a team and what is not feasible, without making any claim which equilibrium is actually chosen (with the exception that we focus on symmetric equilibria). As a first result, we can show that if agents do *not* exhibit inconsistent time preferences, forming a team is not feasible.

[8] If the team cannot be dissolved, for example because the team is an entity within a firm, then agents would play the static equilibrium, yielding an even worse utility than individual production. We discuss this possibility below, in Section 6.

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

Obviously, a team is not needed if $\beta = 1$. We show that forming a team even is not possible in that case. This is driven by two aspects. On the one hand, the standard free-rider problem of team production is present, making an underprovision of effort optimal in the short run. On the other hand, an agent's outside option is already at the first best. Hence, a breakdown of the team is associated with no costs and a deviation always more tempting than working on the joint project. Furthermore, teamwork is only (potentially) feasible for effort levels strictly above e^{I}.

Lemma 2 *No effort level $e^T \leq e^I$ can be enforced within a team.*

The intuition of Lemma 2 is similar to the one driving Lemma 1. For $e^{\mathrm{T}} \leq e^{\mathrm{I}}$, continuation utilities of individual production are higher than those of teamwork. Together with the free-rider problem, this indicates that teamwork is not only not worthwhile, but not even feasible for $e^T \leq e^I$. Lemma 2 also implies that if a team can be formed, this is optimal for both agents; because the associated effort is higher than e^{I}, teamwork can help them to overcome their self-control problems.

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

Proposition 1 *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.*

For δ sufficiently large, today's value of future cooperation becomes so large that it necessarily dominates today's gain of a deviation. Proposition 1 establishes our first main result – that teamwork can help to overcome self-control problems. The next proposition makes the feasibility of teamwork more precise.

Proposition 2 *Positive effort within a team can be enforced if and only if*

$$\delta \geq \underline{\delta} \equiv \frac{2\sqrt{4 - 6\beta + 3\beta^2} - 1}{5 - 8\beta + 4\beta^2}. \tag{8}$$

Furthermore, $d\underline{\delta}/d\beta \geq 0$.

To obtain $\underline{\delta}$, we derive the level of team-effort that maximizes the left-hand-side of the (IC) constraint, denoted $\underline{e}^{\mathrm{T}}$. Since it is unique, teamwork is only feasible if the (IC) constraint holds for $\underline{e}^{\mathrm{T}}$. Two aspects are important. First of all, $\underline{\delta} < 1$ for $\beta < 1$, hence agents with self-control problems can generally form productive teams. Furthermore, $d\underline{\delta}/d\beta \geq 0$ implies that a lower β generally makes it *easier* to enforce any effort within a team.

The latter point is not that straightforward, since a lower β generally has two countervailing effects. On the one hand, it reduces e^{I} and an agent's outside option, thereby relaxing the (IC) constraint. On the other hand, the future becomes less valuable, which tightens the (IC) constraint. However, at the threshold $\underline{\delta}$ only effort $\underline{e}^{\mathrm{T}}$ can just be enforced, which is increasing in β (furthermore, $\underline{e}^{\mathrm{T}} \to 0$ for $\beta \to 0$). Therefore, a lower β implies that the critical threshold of δ above which a team can be formed is reduced, however the enforceable effort at this threshold goes down.

The grey line in Figure 1 below gives $\underline{\delta}$ as a function of β; the shaded region above gives all combinations of δ and β for which positive effort within a team can be enforced.

We are particularly interested in the conditions for which first-best effort $e^{\mathrm{FB}} = \delta V/c$ can be enforced. In this case, the (IC) constraint becomes

$$-\left(1-\beta\left(1-\beta\right)\right)+\delta\left(1-\beta^{2}\left(1-\beta\right)\right)\geq 0. \tag{9}$$

Since the term $\delta^2 V^2/2c$ cancels out, the value of the fundamentals V and c has no effect on the enforceability of team-effort. Only the ratio V^2/c determines the size of the left hand side, however does not affect whether it is positive. This implies

Proposition 3 *First-best effort e^{FB} within a team can be enforced if*

$$\delta \geq \delta^{FB} \equiv \left(1-\beta\left(1-\beta\right)\right)/\left(1-\beta^{2}\left(1-\beta\right)\right). \tag{10}$$

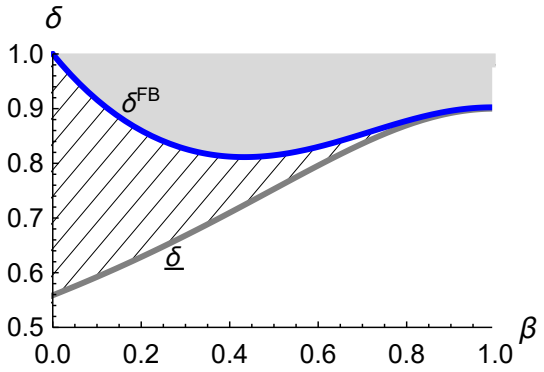
Note that $\delta^{\mathrm{FB}} < 1$ for $\beta \in (0,1)$. Furthermore, δ^{FB} increases in β for large initial values of β and decreases for small initial values of β.[9] Therefore, more severe self-control problems of team members can make it easier to sustain first-best effort

[9] Formally, $d\delta^{\mathrm{FB}}/d\beta = -\left[\left(1-\beta\right)\left(1-3\beta+\beta^{2}\left(1-\beta\right)\right)\right]/\left(1-\beta^{2}\left(1-\beta\right)\right)^{2}$.

within a team. Again, this is driven by the two effects of a lower β on the enforceability of a given effort level e^{T}. A lower β not only amplifies discounting and therefore directly tightens the (IC) constraint, it also relaxes the (IC) constraint by decreasing off-equilibrium individual effort levels in the future (i. e., e^{I} is reduced) and consequently agents' outside options from today's perspective. Starting from $\beta = 1$ and reducing β, the second effect initially dominates if $e^{\mathrm{T}} = e^{\mathrm{FB}}$. For rather low values of β, the first effect dominates.

In the following Figure 1, the blue line gives δ^{FB}, and the grey region above shows all combinations of δ and β for which e^{FB} can be enforced.

Figure 1: Region Where Positive Effort and the First-Best can be Attained



Note: The lower curve gives $\underline{\delta}$ as defined in Proposition 2. Above this curve, positive effort in a team is feasible. The upper curve gives δ^{FB} as defined in Proposition 3. Above this curve, a relational contract can implement the first-best effort level.

Concluding, more severe self-control problems can help to improve the performance of a team (which – if feasible – always yields higher effort than e^{I}, see Lemma 2). This is a general feature of relational contracts, which work better if agents are vulnerable. Someone who is locked in a relationship because their outside option is unattractive is willing to sacrifice more in order to maintain cooperation. An agent's vulnerability might be more pronounced if finding an adequate replacement for one's partner is impossible or – as in our case – if being thrown back on one's own is particularly bad.

4 Extensions

In the following, we extend our setup along two lines and show that a couple of further interesting results can be obtained.[10] First, we assume that teamwork is associated with exogenous technological benefits. Second, we explore the implications of agents not being (fully) aware of their future self-control problems.

4.1 Teamwork with Exogenous Benefits

We have shown that teamwork can help to overcome an agent's self-control problems. For time-consistent agents, teamwork is not possible – however also not needed. Here, we show that even if teamwork renders technological benefits, implying that also time-consistent agents would rather work within a team than on individual projects, time-inconsistent agents can perform better than time-consistent ones. This is true as long as the exogenous benefits of teamwork are not too large. The mechanism driving this result is equivalent to the one underlying our previous analysis: A lower β not only reduces continuation utilities on the equilibrium path, but also agents' off-path utilities. As long as the latter aspect dominates, a lower β can induce a higher performance within the team.

We focus on one particular case of exogenous team benefits and assume that if both agents work on a joint project, the probability to generate the payoff V in period $t+1$ is $e_1 + e_2 + 2\alpha\min\{e_1, e_2\}$, with $\alpha \geq 0$ (and impose the assumption $\delta V(1+\alpha)/c < 1/2$ to always guarantee an interior solution). We use the minimum-function in order to make sure that the exogenous benefits of teamwork are only realized if *both* actually work on the joint project. Other specifications for the exogenous benefits of teamwork would generate very similar results. A value $\alpha = 0$ yields the situation analyzed above; a value $\alpha > 0$ could be generated by discussions of the team members about the joint problems which deepens each agent's understanding, or by heterogeneities in the agents' abilities to tackle different aspects of a project. If both agents exert team effort e^{T}, the probability to generate the payoff V in period $t+1$ is $2e^{\mathrm{T}}(1+\alpha)$.

Before analyzing the feasibility of teamwork, we have to be precise about the defi-

[10] Further extensions can be found in an Online Appendix.

nition of first-best effort in this section. First-best effort – as regarded from earlier periods – now is different under individual production than within a team. Here, we focus on the highest feasible payoff an agent can possibly expect, which implies that the technological benefits of teamwork are enjoyed. Hence, we define first-best effort levels e_1^{FB} and e_2^{FB} as maximizing the joint team payoff as regarded from earlier periods, i. e.

$$-\frac{ce_1^2}{2} - \frac{ce_2^2}{2} + \delta\left(e_1 + e_2 + 2\alpha\min\{e_1, e_2\}\right)V. \tag{11}$$

Since a potential output V is shared equally, no other definition of first-best effort could make both agents better off. The symmetric first-best effort level e^{FB} is thus

$$e^{\mathrm{FB}} = \frac{\delta\left(1+\alpha\right)V}{c}. \tag{12}$$

Furthermore, an agent's equilibrium utility stream given both agents exert team effort e^T is

$$U^{\mathrm{T}} = -c\frac{(e^{\mathrm{T}})^2}{2} + \beta\delta e^{\mathrm{T}}(1+\alpha)V + \beta\frac{\delta}{1-\delta}\left(\delta e^{\mathrm{T}}(1+\alpha)V - c\frac{(e^{\mathrm{T}})^2}{2}\right). \tag{13}$$

Individual production still constitutes agents' outside options[11] and is not affected by the existence of technological team benefits. It yields a per-period utility $u^{\mathrm{I}} = -c\,(e^{\mathrm{I}})^2/2 + \beta\delta e^{\mathrm{I}}V$, and optimal effort for each agent is $e^{\mathrm{I}} = \beta\delta V/c$.

An agent's deviation utility is hence given by

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

and the (IC) constraint boils down to

$$\begin{aligned}\left(\beta\delta e^{\mathrm{T}}\left(\frac{1}{2}+\alpha\right)V - c\frac{(e^{\mathrm{T}})^2}{2}\right) - \left(\beta\delta e^{\mathrm{I}}V - c\frac{(e^{\mathrm{I}})^2}{2}\right)\\ +\beta\frac{\delta}{1-\delta}\left[\left(\delta e^{\mathrm{T}}\left(1+\alpha\right)V - c\frac{(e^{\mathrm{T}})^2}{2}\right) - \left(\delta e^{\mathrm{I}}V - c\frac{(e^{\mathrm{I}})^2}{2}\right)\right] \geq 0. \qquad \text{(IC')}\end{aligned}$$

[11]Note that, if the exogenous team benefits are large enough (more precisely, if $\alpha \geq 1$), then there also are static Nash equilibria where agents team up. The largest feasible punishment can be achieved (and hence maximizes equilibrium utilities, see Abreu (1988)), though, if any deviation is followed by a reversion to the static Nash equilibrium with lowest utility levels, that is, individual production.

Generally, a larger α helps to enforce cooperation within a team, irrespective of whether agents have a present bias or not. Hence, a larger α lets potential extra benefits of a lower β diminish. As long as α is not too large, though, the performance of teams with time-inconsistent agents can be better than of teams without inconsistencies. To see that, we focus on first-best effort e^{FB} and the conditions under which it can be enforced. For e^{FB}, the (IC') constraint becomes

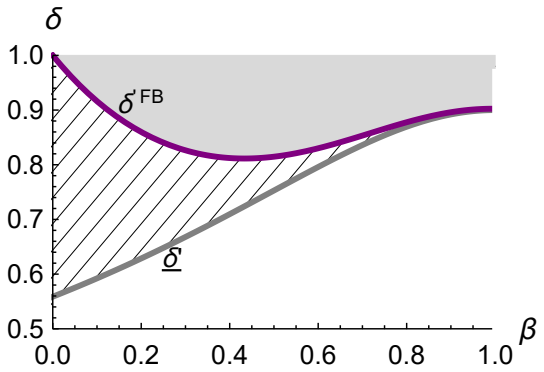
$$-(1+\alpha)^2-\beta^2+\beta(1+\alpha)(2\alpha+1)+\delta\left((1+\alpha)^2-\beta(1+\alpha)\alpha-\beta^2(1-\beta)\right)\geq 0,$$

and first-best effort is feasible for

$$\delta \geq \delta'^{FB} \equiv \frac{(1+\alpha)^2+\beta^2-\beta(1+\alpha)(2\alpha+1)}{\left[(1+\alpha)^2-\alpha\beta(1+\alpha)-\beta^2(1-\beta)\right]}. \tag{14}$$

As an example, assume that $\alpha = 0.05$, i. e. teamwork boosts total productivity by 5 % compared to individual production. Then, the blue line in the following Figure 2 gives δ'^{FB} as a function of β; the grey region above gives all combinations of δ and β for which e^{FB} can be enforced for $\alpha = 0.05$. Furthermore, for the sake of comparability, the grey line gives $\underline{\delta'}$ – the threshold above which any positive effort can be enforced in a team – and the shaded region above all combinations of δ and β for which this is the case for $\alpha = 0.05$.[12]

Figure 2: Region Where Positive Effort/the First-Best can be Attained, $\alpha = 0.05$



Note: The lower curve gives $\underline{\delta'}$ for $\alpha = 0.05$. Above this curve, positive effort in a team is feasible. The upper curve gives δ'^{FB} as defined in Equation (14). Above this curve, a relational contract can implement the first-best effort level.

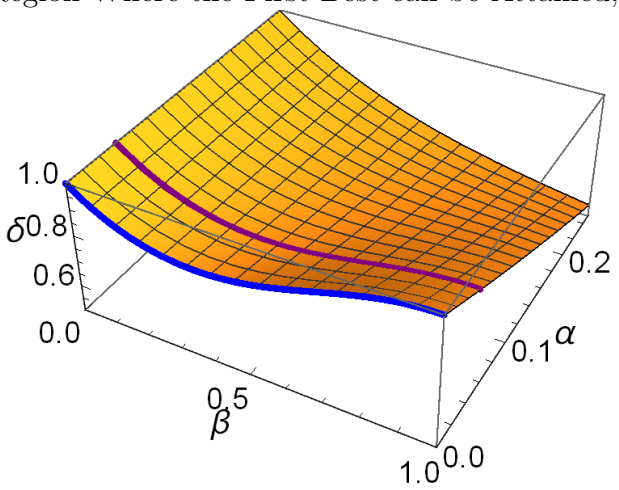
Hence, there exist values of δ such that first-best effort cannot be enforced when

[12] For general values α, positive team-effort can be implemented if $\left(\frac{1}{2}(1+\delta)+\alpha\right)^2 - \left(1-\delta^2(1-\beta)^2\right) \geq 0$.

$\beta = 1$, but that this is feasible for values $\beta < 1$.

Concluding, even in the presence of exogenous technological benefits of teamwork, teams of agents with self-control problems can perform better than teams of agents without – however only if α is not too large. The following Figure 3 gives δ'^{FB} for general values of α. The region above the curve shows all combinations of δ, β and α for which e^{FB} can be enforced.

Figure 3: Region Where the First-Best can be Attained, General α



Note: Above the surface, the first-best can be attained. The blue curve stands for $\alpha = 0$, hence it coincides with the blue curve of Figure 1. The purple curve stands for $\alpha = 0.05$, hence it coincides with the purple curve of Figure 2. The slope with respect to α is negative, which means that with positive technological benefits from teamwork, it is easier to reach the first-best with the help of a relational contract.

4.2 Naive and Partially Naive Agents

The agents in our setup are sophisticated in a sense that they can perfectly anticipate their future self-control problems and hence their 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

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

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 (Lemma 1): 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.

5 Discussion

In this section, we assess the general validity of our results. First, we discuss whether an agent should be able to play a relational contract not only with another individual, but also with his future selves. Second, we compare our approach to other potential mechanisms to enforce cooperation within a team, and discuss how those could be differentiated empirically.

5.1 Intra-personal Game

From a formal point of view, the relational contract we propose could also be formed by an individual agent playing a game with his future selves, conditioning current actions on past behavior. Such games have been analyzed by Laibson (1994) or Bernheim, Ray, and Yeltekin (2013), who allow for "personal rules" which are formally characterized by subgame perfect equilibria of the game played between different selves (note that this is different from widely-analyzed situations where the current self restricts a future self's choice set, which is not feasible in our setting). However, we think that Markov perfect equilibria have a larger intuitive appeal when analyzing intra-personal decision making (like in Krusell and Anthony A. Smith (2003) or Basu (2011)). In the following, we also make two more substantive arguments to support our view that relational contracts in teams are better suited to tackle self-control problems than intra-personal rules that require self-inflicted punishment. First, we present a formal argument for why renegotiation can better be prevented in teams than in an arrangement an agent has with himself. Second, we argue that agents are more likely to punish deviations of other agents than deviations of themselves – because in the latter case, they too easily accept excuses for their own behavior and because deviations are more salient.

5.1.1 Renegotiation

Due to inconsistent time preferences, players would *at a given point in time* benefit from renegotiating their agreement and postpone high effort to the next period. This kind of renegotiation would be deemed optimal in every period, hence potentially make cooperation within a team or an accordingly designed intra-personal game entirely infeasible. We argue that a team can do better to prevent such renegotiation than an individual can do himself.

First, the relational contract might be adjusted accordingly: At the beginning of the game, both agree that any attempt to renegotiate the agreement later on is regarded a deviation and triggers a separation. Although this aspect can of course also be subject to renegotiation, we would argue that renegotiation attempts shape expectations that this will happen over and over again. But cooperation in a team only works of no one expects a renegotiation of the agreement in the future, therefore individuals might abstain from renegotiation in order to not affect expectations on

future behavior. An individual, on the other hand, can arbitrarily adjust his plans at any point in time and do what currently seems optimal (as we argue in the next section, a deviation in an intra-personal game is less salient than within a team). Put differently, being in a team implies that any action implies a reaction – and this reaction depends on whether behavior has previously been agreed-upon or not.

Second, we can make a more formal argument and characterize an equilibrium that prevents any renegotiation attempts even if those are not punished by a surplus-reducing permanent reversion to individual production. We provide a full characterization of this equilibrium in the Online Appendix, here we just sketch its properties and argue that it does not work for intra-personal games. To establish an equilibrium that prevents on-path renegotiation (that is, renegotiation of the initial agreement; renegotiation of an off-path punishment is also analyzed in the respective section of the Online Appendix), we have to adjust the model setup in one dimension: The outside option is not necessarily individual production anymore; instead, players are able to find new partners with whom they can form a team.

Equilibrium play in a new team is designed in a way that for a given team, high effort can be enforced in every period and on-path renegotiation does not occur. The equilibrium has the properties that renegotiation triggers a termination in the *subsequent* period. This termination threat must be costly from today's perspective in order to prevent renegotiation attempts. Furthermore, it must be optimal from tomorrow's perspective if renegotiation has occurred, but not otherwise. This latter aspect implies that a termination does not make any agent worse off at the time when this decision is made and hence is renegotiation-proof. Therefore, at the beginning of a period, agents have to be indifferent between staying in their current relationship and going for a new match. All this is obtained by letting new matches start with a number of periods with individual production and move to high effort thereafter. The number of periods with individual production is found by equalizing an agent's utility of remaining in his current match (and continuing with high effort) with an agent's utility of a separation (also taking into account that he might not immediately find a new match). Furthermore, an agent's inconsistent time preferences imply that from today's perspective, he strictly prefers high team effort from tomorrow on. Therefore, such an equilibrium can prevent attempts for a renegotiation of the initial agreement. Furthermore, we assume that equilibrium play in new matches can also be contingent on past free-riding in order to prevent the latter. This requires some transparency in a sense that prospective new partners

are aware of an agent's past deviations.

Such an equilibrium that prevents on-path renegotiation can not be sustained if everyone plays an intra-personal game to sustain high effort because it requires the opportunity to go for a new relationship. In an intra-personal game, an individual rather has the opportunity to renegotiate any agreement with himself, leading to a constant delay of high effort.

5.1.2 Punishment is More Likely in Teams

In intra- as well as interpersonal arrangements, effort above e^I can only be implemented if deviations are deterred. Those can take the form of free-riding, but also – as discussed in the previous section – of renegotiation attempts to postpone delivery of high effort. Any deviation must be regarded a violation of the agreement and be punished adequately (where a punishment can also mean that a team breaks up and its members go for new matches, as analyzed in the previous section). Whereas identifying a deviation is straightforward in our model, the situation is less clear in many real-word situations. In particular, individuals sometimes do have valid excuses for postponing high effort – may it be for private reasons or because of other jobs that have to be done – and a punishment only seems justified if an excuse is *not* regarded valid. However, the validity of excuses can be subject to interpretation.[13] In psychology, for example, there is evidence that a "self-serving bias" or "attribution bias" makes individuals prone to generally accept their own excuses. Such a bias has been explained by Zuckerman (1979), who reports the notion that individuals "attempt to enhance or protect their self-esteem by taking credit for success and denying responsibility for failure" (p. 245). Moreover, individuals attribute positive outcomes to inherent personality traits, whereas negative outcomes are caused by situational influences (see Moskowitz (2005), or Forsyth (2008) among many others).[14]

[13] Although these aspects have a flavor of games with incomplete or asymmetric information (which we analyze in the Online Appendix), a potential ambiguity when assessing the validity of an excuse addresses a different dimension.

[14] As examples for empirical evidence, take Lau and Russell (1980), who show that members of sports teams are much more likely to attribute success internally than losses; or Stewart (2005), who analyzes a survey completed by survivors of motor vehicle crashes. In particular people who were involved in more severe crashes put responsibility on other drivers or road/weather conditions. In economics, such an attribution bias has only received limited attention. An exception is Hestermann and Le Yaouanq (2016), who show that overconfidence concerning one's own ability

In our case, such a self-serving bias would imply that whenever someone realizes that they have not worked as hard as intended, they rather blame factors outside their own responsibility – and therefore do not see a need to "punish" themselves after a deviation. If this is (subliminally) anticipated by individuals, performance is negatively affected.

The situation is different when someone else's behavior is assessed, which has been regarded as an "actor-observer bias". Jones and Nisbett (1987, page 80), for example, provide the following definition:[15] "The actor's view of his behavior emphasizes the role of environmental conditions at the moment of action. The observer's view emphasizes the causal role of stable dispositional properties of the actor. We wish to argue that there is a pervasive tendency for actors to attribute their actions to situational requirements, whereas observers tend to attribute the same actions to stable personal dispositions."

Therefore, teamwork can also act as a commitment device to not bring up an excuse if one did not work hard, because one knows that their counterpart will not accept the excuse and punish them accordingly.

Finally, we argue that deviations in a team are more salient than deviations in an intra-personal game: In the former case, a deviation either means one agent free-riding on the other agent's effort (implying a clear violation of the agreement), or a renegotiation of the relational contract. Then, one agent effectively has to approach the other agent to suggest postponing high effort. In an intra-personal game, both kinds of deviations just boil down to exerting a different effort level than intended, which can rather be shrugged off by individuals. Therefore, deviations are more salient in a team, which we argue implies that those are more likely to be punished in real-world situations.

5.2 Comparisons and Predictions

We have shown that even if teamwork renders no technological benefits, the associated relational contract can help overcome problems of self control. Naturally, we do

can lead to an attribution bias.

[15] For further explorations of the actor-observer bias see Moskowitz (2005), Forsyth (2008), who also explore the actor-observer bias and provide explanations for its foundation.

not claim that our mechanism is the sole reason for the formation and functioning of teams. Still, in order to assess its real-word relevance, it is important to lay out how our approach can be distinguished from other potential drivers.

In this section, we analyze two other potential forces that might increase effort in teams – peer pressure and reciprocity – and discuss to what extent those would yield different predictions (that could in principle be tested in lab or field experiments) than ours. We focus on these two because they appear rather related. Other reasons for an endogenous formation of teams, for example the existence of multitasking problems (Corts (2007)), work under entirely different settings, and the differences to our concept appear evident.

Before relating peer pressure and reciprocity to our approach, let us first be more precise about their meaning. With both, preferences are assumed to entail "social" elements that foster cooperation. Peer pressure as a means to address free-rider problems has (at least in economics) been introduced by Kandel and Lazear (1992). They assume that each individual's utility function also contains a peer pressure term that depends on one's own as well as on others' actions. Effort reduces peer pressure, which creates additional incentives to exert effort. Peer pressure works via guilt or shame: Guilt is an internal mechanism and does not require mutual monitoring, whereas shame can only emerge once free riding is detected by others.

Preferences for reciprocity (see Akerlof (1982), Falk and Fischbacher (2006), or Englmaier and Leider (2012)) make individuals also care about their peers' utilities – provided those have acted to one's own benefit. In a team, effort has a public goods component. Therefore, an agent's effort might increase the other's utility – who then wants to reciprocate by working harder, and vice versa.

Now, when comparing various approaches, one has to differentiate between a *rationale* for team formation and the *mechanism* to enforce high effort within a team and overcome the free-rider problem. In our setting, both aspects are relevant: The rationale for forming a team is to overcome problems of self control. Given it has been formed, effort is enforced by repeated-game incentives, that is, a relational contract between team members (see Che and Yoo, 2001). But also the very existence of self-control problems helps to enforce team effort – because those reduce each players' outside option in case punishment means a reversion to individual production. This integral perspective on the formation and functioning of teams is not shared by models of peer pressure or reciprocity (or other concepts we are aware of). These

mechanisms are supposed to help enforce high effort, taking the existence of the team as granted.

Given a team has been formed, our concept differs from peer pressure and reciprocity in a sense that both also work in a static setting. One might therefore compare team effort in a static setting with team effort in a setting with repeated interaction. Ceteris paribus higher effort in the second case would indicate that relational contracts indeed help to overcome free-rider problems. Landeo and Spier (2015) provide experimental evidence for this conjecture. They let agents work in teams and show that repeated interaction and mutual monitoring ceteris paribus makes them work harder.

Furthermore, even if we assume that individuals suffer from self-control problems, and if they have the option to form a team, outcomes might differ depending on whether the mechanism to enforce high effort is a relational contract, or whether it is peer pressure or reciprocity:

Whereas our mechanism always makes it optimal to form a team (given this is feasible; see the discussion following Lemma 2), this does not necessarily hold with peer pressure. Peer pressure produces higher effort levels, but the pressure itself is a cost that has to be borne by team members. Therefore, if peer pressure is high in a team, agents might rather work on individual projects with lower effort – because they feel badly about being in an arrangement with substantial peer pressure (see Lazear and Shaw (2007)).

The situation is less straightforward with reciprocity, for the following reason: Reciprocal preferences create additional utility if effort is above a certain benchmark (a "warm-glow" effect as in Akerlof (1982) or Englmaier and Leider (2012)), or negative utility if effort is below such a threshold (see Falk and Fischbacher (2006)).

As an example (which is similar to Falk and Fischbacher (2006) and Englmaier and Leider (2012)), assume that agent 1's utility in a team is $u_1^T = \beta\delta\frac{e_1^T+e_2^T}{2}V - \frac{c(e_1)^2}{2} + \eta e_1^T\left(e_2^T - \overline{e}\right)$, where $\eta \geq 0$ is a reciprocity parameter and $\overline{e}$ the benchmark effort level.[16] If $\overline{e}$ is rather low (and η sufficiently high), team-effort would easily be above the benchmark, induce effort levels above e^I and furthermore create additional

[16]To be fully precise, u_1^T is defined as a function of agent 1's beliefs about agent 2's team effort (since effort choices are made simultaneously), which however coincide with agent 2's exerted effort on the equilibrium path.

utility via the reciprocity term. In this case, forming a team might even be optimal for time-consistent agents in order to enjoy the associated warm glow – a result that could not be generated by our model (see Lemma 1). If $\overline{e}$ is rather high, on the other hand, the situation would be rather similar to a model of peer pressure. Then, if e^T is below the benchmark, having a team is associated with a utility loss. If this utility loss is too high in relation to the increase in effort levels, agents rather work on individual projects.

As a final difference that can potentially be assessed empirically, recall that the performance of a relational contract generally decreases with an agent's outside option. Therefore, more severe self-control problems might actually improve team performance – an outcome that could not be generated with peer pressure or reciprocity.

6 Conclusion

We have shown that teamwork can serve as an implicit commitment device to overcome problems of procrastination and self-control. Even if teamwork renders technological benefits, the team-performance of agents with self-control problems can actually be better than the performance of agents without. We have focussed on a setting with just two individuals and abstracted from standard employment situations where a principal assigns agents to jobs, and in particular has the option to let them work individually or form a team. In such an environment, Che and Yoo (2001) have shown that it can be optimal for a principal to let agents (without self-control problems) work in a team – if agents interact repeatedly and can therefore form a relational contract, if the principal can only use an imperfect performance measure to compensate individuals, and if agents can mutually monitor each other's effort. Furthermore, agents are protected by limited liability, hence the principal cannot extract high rents ex ante that she has to grant agents when incentivizing them.

Naturally, the results derived in Che and Yoo (2001) also hold if agents have self-control problems. In this case, the benefits of a team will be even larger than if agents do not suffer from a present bias. Generally, mutual monitoring allows to implement higher effort levels within a team if the standard discount factor δ is sufficiently large. Then, provided the principal's return to effort is larger than the

agents' return V (which in this case might represent a bonus agents receive for a high output), she will assign agents to teams rather than letting them work on individual projects. There, note that enforcing high effort actually becomes easier with a principal than in the setting we have analyzed in the present paper: If agents have to stay in the team after a deviation, their off-path utilities are reduced, which reduces their temptation to deviate. The exact benefits of teamwork in a setting with principal is present will however also rely on further aspects – like whether a principal can commit to pay an output-based bonus, or whether agents are free to leave in every period (in particular after a deviation). We think that it is worthwhile to study these aspects in future research, and more generally address the question of optimal workplace design for employees with self-control problems.

A Appendix – Proofs

Proof of Lemma 1. Note that in this case, $e^{\mathrm{I}} = e^{\mathrm{FB}}$; the (IC) constraint becomes

$$\left(\delta e^{\mathrm{T}}\frac{V}{2} - \frac{c\,(e^{\mathrm{T}})^2}{2}\right) - \left(\delta e^{\mathrm{FB}}V - \frac{c\,(e^{\mathrm{FB}})^2}{2}\right)$$
$$+\frac{\delta}{1-\delta}\Big[\big(\delta e^{\mathrm{T}}V - \frac{c\,(e^{\mathrm{T}})^2}{2}\big) - \big(\delta e^{\mathrm{FB}}V - \frac{c\,(e^{\mathrm{FB}})^2}{2}\big)\Big] \geq 0 \qquad (15)$$

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

Proof of Lemma 2. The proof is almost equivalent to that of Lemma 1. Assume that $e^{\mathrm{T}} \leq e^{\mathrm{I}}$. Because $e^{\mathrm{I}} \leq e^{\mathrm{FB}}$ and $\delta eV - \frac{ce^2}{2}$ is increasing for effort levels below e^{FB}, the second line of the (IC) constraint, $\big(\delta e^{\mathrm{T}}V - c\,(e^{\mathrm{T}})^2/2\big) - \big(\delta e^{\mathrm{I}}V - c\,(e^{\mathrm{I}})^2/2\big) \leq 0$ for $e^{\mathrm{T}} \leq e^{\mathrm{I}}$; because e^{I} maximizes $\beta\delta eV - ce^2/2$ the first line of the (IC) constraint, $\beta\delta e^{\mathrm{T}}V/2 - c\,(e^{\mathrm{T}})^2/2 - \big(\beta\delta e^{\mathrm{I}}V - c\,(e^{\mathrm{I}})^2/2\big)$, is strictly negative. Therefore, the left hand side of the (IC) constraint is strictly negative for $e^{\mathrm{T}} \leq e^{\mathrm{I}}$. ■

Proof of Proposition 1. The second line of the (IC) constraint,

$$\big(\delta e^{\mathrm{T}}V - c\,(e^{\mathrm{T}})^2/2\big) - \big(\delta e^{\mathrm{I}}V - c\,(e^{\mathrm{I}})^2/2\big),$$

is strictly positive for any $\beta < 1$ and $e^{\mathrm{I}} < e^{\mathrm{T}} \leq e^{\mathrm{FB}}$. Following Lemmas 1 and 2, the first line of the (IC) constraint is negative, however it is bounded. Hence, for $\delta \to 1$, the (IC) constraint is satisfied for any e^{T} with $e^{\mathrm{I}} < e^{\mathrm{T}} \leq e^{\mathrm{FB}}$. ■

Proof of Proposition 2. First, we derive the level of e^{T} that maximizes the left-hand-side of the (IC) constraint. Only if (IC) holds for this effort level, positive effort within a team can at all be enforced. The first derivative of the left-hand-side of (IC) with respect to e^{T} is

$$\beta\delta\frac{V}{2} - ce^{\mathrm{T}} + \frac{\beta\delta}{1-\delta}\left(\delta V - ce^{\mathrm{T}}\right),$$

hence the left-hand-side of (IC) is maximized for $\underline{e}^{\mathrm{T}} = \frac{\beta\delta\frac{V}{2}(1+\delta)}{c(1-\delta+\beta\delta)}$ (the second-order condition holds since the second derivative of the left-hand-side of (IC) with respect to e^{T} equals $-c\frac{1-\delta+\beta\delta}{1-\delta} < 0$). Plugging this value into (IC) and re-arranging gives

$$-3 + 2\delta + 5\delta^2 + 4\beta\delta^2\left(\beta - 2\right) \geq 0. \tag{16}$$

Solving for δ yields $\underline{\delta} = \left(-1 + 2\sqrt{4 - 6\beta + 3\beta^2}\right)/\left(5 - 8\beta + 4\beta^2\right)$. Finally,

$$\frac{d\underline{\delta}}{d\beta} = 2\left(1-\beta\right)\frac{17 - 24\beta + 12\beta^2 - 4\sqrt{4 - 6\beta + 3\beta^2}}{\sqrt{4 - 6\beta + 3\beta^2}\left(5 - 8\beta + 4\beta^2\right)^2}. \tag{17}$$

The sign of $d\underline{\delta}/d\beta$ is determined by the sign of its numerator, where

$$\begin{aligned} 17 - 24\beta + 12\beta^2 - 4\sqrt{4 - 6\beta + 3\beta^2} &= 1 + 4\left(4 - 6\beta + 3\beta^2\right) - 4\sqrt{4 - 6\beta + 3\beta^2} \\ &= 1 + 4\Big[\left(3\left(1-\beta\right)^2 + 1\right) - \sqrt{3(1-\beta)^2 + 1}\Big]. \end{aligned}$$

Since $3\left(1-\beta\right)^2 + 1 \geq 1$, we have $\Big[\left(3\left(1-\beta\right)^2 + 1\right) - \sqrt{3(1-\beta)^2 + 1}\Big] \geq 0$. Thus, the numerator is positive. ■

B Online Appendix

B.1 Renegotiation

As with many cases of subgame-perfect equilibria in repeated games, punishing free-riding by permanently switching to individual production afterwards is not renegotiation-proof because both agents would benefit from re-starting the team. This, however, would implicate adverse effects on the ex-ante incentives of anticipating agents. Still, such surplus-destroying off-path punishment could be supported by the following, slightly informal, argument. An agent is only willing to contribute to the team if he expects the other agent to do so as well. A deviation of one agent then permanently deteriorates the other agent's expectations that he will contribute in the future. Consequently, any attempt for renegotiation following a deviation would not help to regain cooperation – because both agents do not expect their counterpart to cooperate anymore. But even if we require punishments to be renegotiation-proof, we show below that teamwork can nevertheless be sustained.

Moreover, there also is an opportunity for mutually beneficial renegotiation *on* the equilibrium path in our setting. In any period t, both players would be better off *at this point in time* by renegotiating their agreement, work on individual projects today and resume teamwork from the next period on. Since renegotiation would be deemed optimal in every period, it would potentially make cooperation within a team entirely infeasible.

This dilemma is caused by inconsistent time preferences. We argue that teamwork is a mechanism to overcome the associated commitment problem, for example by adjusting the relational contract accordingly: At the beginning of the game, both agree that any attempt to renegotiate the agreement later on is regarded a deviation and triggers a separation. Of course, this part of the agreement can also be subject to renegotiation. But then, we can refer to the above-mentioned argument using expectations on future behavior: Cooperation in a team only works if no one expects a renegotiation of the agreement in the future. Being approached by one's partner to renegotiate the agreement (and agreeing to it) could then induce expectations that this will happen over and over again – and prevent renegotiation in the first place.

We think that this argument has a lot of intuitive appeal. However, one might

also argue that initiating a renegotiation is fundamentally different from free-riding. Whereas the latter is a clear deviation of the agreement at the expense of one's counterpart, renegotiation can result in a jointly beneficial change of the contractual terms. Therefore, one might claim that renegotiation should not trigger the same kind of punishment as free-riding does. But then, agents would have incentives to repeatedly renegotiate their agreement and consequently not be able to sustain higher effort. In the following, we show that if an attempt for renegotiation is *not* considered a deviation that is supposed to be punished by a surplus-reducing permanent reversion to individual production, high effort within a team can nevertheless be sustained if our setup is slightly extended: The outside option is not necessarily individual production anymore; instead, players are potentially able to find new partners with whom they can form a team. More precisely, assume that at the beginning of every period, players decide whether they stay in their current match, or whether they leave and look for a new one. We do not fully model the matching market, but assume that a new match can be found with probability $\gamma \leq 1$.[17] Before designing an equilibrium that can prevent renegotiation, note that (for now) we still let a free-rider be punished by having to rely on individual production thereafter. This requires some transparency in a sense that prospective new partners are aware of an agent's past deviations, which we assume. Below, we also analyze renegotiation-proof punishments that do not require a complete reversion to individual production.

Now, we construct an equilibrium that implements high team effort in every period and therefore prevents (attempts for) on-path renegotiation. In this equilibrium, we want renegotiation to trigger a termination in the following period. Second, a termination in the following period has to be costly from today's perspective. But such a termination threat must be credible (and not be subject to renegotiation as well). Therefore, the third property of this equilibrium is that, when tomorrow comes, a termination must be optimal if renegotiation has occurred in the previous period – but not otherwise. This implies that a termination does not make any agent worse off at the time when this decision is made and hence is renegotiation-proof. The following equilibrium has all these properties. Any new match starts with a number of periods with individual production and moves to high effort thereafter. The number of periods with individual production is obtained by equalizing an agent's utility of

[17]For example, this could be endogenized by assuming that existing matches break up for exogenous reasons with some probability, and hence there always exist potential new partners.

remaining in his current match (and continuing with high effort) with an agent's utility of a separation (taking into account that he might not immediately find a new match). If γ is sufficiently large, such an equilibrium exists. Furthermore, because of an agent's self-control problems, tomorrow's indifference between an equilibrium with continuously high team effort, and a combination of individual production and teamwork, implies that from today's perspective, he strictly prefers high team-effort from tomorrow on. Therefore, such an equilibrium can prevent attempts for renegotiation. Note that renegotiation to equilibria which are not maximizing the surplus can also be prevented. For example, a team might renegotiate and agree that individual production is chosen for two periods, rendering a separation suboptimal also from tomorrow's perspective. This, however, could be dealt with by assuming that a termination happens in the period where agents are supposed to resume team production.

In the following proposition, we provide a lower bound on γ above which first-best effort e^{FB} in such a renegotiation-proof equilibrium can be sustained provided $\delta \geq \delta^{\mathrm{FB}}$, where δ^{FB} is the discount factor above which first-best effort can be implemented in our main part (see Proposition 3). There, we assume $-1+\delta\left(1+\beta\right)>0$, which is equivalent to having $U^T > U^I$ for $e^T = e^{\mathrm{FB}}$

Proposition 4 *Assume $-1+\delta\left(1+\beta\right)>0$. Allowing for on-path renegotation and provided $\delta \geq \delta^{FB}$, first-best effort e^{FB} within a team can be enforced if*

$$\gamma \geq \overline{\gamma} = \frac{-1+\delta\left(1+\beta\right)+\beta\left(1-\delta^2\right)}{\beta\delta\left(2-\delta\right)}, \tag{18}$$

with $\overline{\gamma} < 1$.

Proof of Proposition 4. We are interested in an equilibrium where team-effort $e^T = e^{\mathrm{FB}} = \delta V/c$ can be implemented in every period, which implies that renegotiation does not happen on the equilibrium path. We first design new matches in a sense that agents are indifferent between staying and going for a new match. When staying and choosing an effort e^{FB}, an agent's utility equals

$$U^T = \beta\delta e^T V - \frac{c(e^T)^2}{2} + \beta\frac{\delta}{1-\delta}\left[\delta e^T V - \frac{c(e^T)^2}{2}\right] = \frac{\delta^2 V^2}{2c}\left[2\beta - 1 + \frac{\beta\delta}{1-\delta}\right].$$

In a new relationship, both agents work on their own projects for $T > 1$ periods. In period $T+1$, both return to teamwork with probability $1-\rho$ and work for another period on their own with probability ρ. From period $T+2$ onwards, they return to teamwork forever after (note that any attempt to renegotiate this agreement is also followed by a termination in the subsequent period). Therefore, the utility of starting a new relationship after having found a match is

$$U^{new} = \beta\delta e^I V - \frac{c\left(e^I\right)^2}{2} + \delta\beta\left\{\left(\delta e^I V - \frac{c\left(e^I\right)^2}{2}\right)\frac{1-\delta^T}{1-\delta} + \delta^T\left[\rho\left(\delta e^I V - \frac{c\left(e^I\right)^2}{2}\right) + (1-\rho)\left(\delta e^T V - \frac{c\left(e^T\right)^2}{2}\right)\right] + \frac{\delta^{T+1}}{1-\delta}\left(\delta e^T V - \frac{c\left(e^T\right)^2}{2}\right)\right\}.$$

For first-best effort, it becomes

$$U^{new} = \frac{\beta^2\delta^2 V^2}{2c} + \delta\beta\frac{\delta^2 V^2}{2c}\left[\frac{(2-\beta)\beta + \delta^T(1-\beta)^2 - \delta^T\rho(1-\beta)^2(1-\delta)}{1-\delta}\right]$$

Given a new match is found with probability $\gamma < 1$, the utility of an unmatched agent equals

$$\tilde{U} = \gamma U^{new} + (1-\gamma)\left(\beta\delta e^I V - \frac{c\left(e^I\right)^2}{2} + \beta\delta\tilde{U}\right),$$

which becomes

$$\tilde{U} = \frac{1}{\left[1-(1-\gamma)\beta\delta\right]}\left\{\frac{\beta^2\delta^2 V^2}{2c} + \gamma\delta\beta\frac{\delta^2 V^2}{2c}\left[\frac{(2-\beta)\beta + \delta^T(1-\beta)^2 - \delta^T\rho(1-\beta)^2(1-\delta)}{1-\delta}\right]\right\}.$$

Now, ρ and T are obtained by setting $U^T = \tilde{U}$, which yields

$$-1 + \delta(1+\beta) + \beta\left(1-\delta^2\right) = \gamma\beta\left[\delta(\beta + (1-\delta)) + (1-\beta)\delta^{T+1}(1-\rho(1-\delta))\right].$$

Note that, because of $\delta \geq \delta^{\text{FB}}$, the left-hand-side is positive. It remains to show that there exist values of ρ and T such that this condition holds. There, note that the right-hand-side is maximized for $T = \rho = 0$. These values therefore give us the smallest feasible γ such that this condition holds. For any larger $\gamma \leq 1$, T and ρ can be adjusted accordingly. Hence, there are ρ and T for which this condition holds, if

$$\gamma \geq \overline{\gamma} = \frac{-1 + \delta(1+\beta) + \beta(1-\delta^2)}{\beta\delta(2-\delta)}.$$

Furthermore, $\overline{\gamma} < 1 \Leftrightarrow (1-\delta)(1-\beta) > 0$. Given there is no renegotiation, the (IC) constraint for first-best effort is the same as in our benchmark case, hence yields $\delta \geq \delta^{\text{FB}}$ which holds by assumption. There, recall that we assume that any agent caught free riding cannot be part of a profitable team anymore and is hence forced to stick to individual production afterwards.

It remains to be shown that a renegotiation is never optimal. There, we distinguish between a successful and an unsuccessful renegotiation, where in the former case both work on individual projects for one period (before splitting up and going for new matches). In the latter case, the renegotiation attempt is turned down and both stick to teamwork (before splitting up and going for new matches). There, it is sufficient to show that this holds for $\gamma = 1$. After a successful renegotiation, an agent takes into account that the match is terminated in the subsequent period. Therefore, a successful renegotiation yields utility

$$\begin{aligned} U_S^R &= \beta\delta e^I V - \frac{c\,(e^I)^2}{2} + \beta\delta \left[\left(\delta e^I V - \frac{c\,(e^I)^2}{2} \right) \frac{1-\delta^{T+1}}{1-\delta} \right. \\ &\quad \left. + \delta^{T+1} \left[\rho \left(\delta e^I V - \frac{c\,(e^I)^2}{2} \right) + (1-\rho) \left(\delta e^T V - \frac{c\,(e^T)^2}{2} \right) \right] + \frac{\delta^{T+2}}{1-\delta} \left(\delta e^T V - \frac{c\,(e^T)^2}{2} \right) \right] \\ &= \frac{\beta^2\delta^2 V^2}{2c} + \frac{\beta\delta}{(1-\delta)} \left[\frac{\beta\delta^2 V^2}{2c} (2-\beta) + \delta^{T+1} \frac{\delta^2 V^2}{2c} (1-\rho(1-\delta)) (1-\beta)^2 \right]. \end{aligned}$$

Using $-1 + \delta(1+\beta) = \beta\delta^{T+1}(1-\rho(1-\delta))$ (which provides indifference between staying in and leaving a relationship at the beginning of a period),

$$U_S^R = \frac{\beta^2\delta^2 V^2}{2c} \left(\frac{1+\delta(1-\beta)}{(1-\delta)} \right) + \frac{\delta}{(1-\delta)} \frac{\delta^2 V^2}{2c} \left[-1 + \delta(1+\beta) \right] (1-\beta)^2.$$

This is smaller than U^T if $-1 + \delta(1+\beta) \geq 0$, which holds given $\delta \geq \delta^{\text{FB}}$.

The utility after an unsuccessful renegotiation is,

$$\begin{aligned} U_U^R &= \beta\delta e^T V - \frac{c\,(e^T)^2}{2} + \beta\delta \left[\left(\delta e^I V - \frac{c\,(e^I)^2}{2} \right) \frac{1-\delta^{T+1}}{1-\delta} \right. \\ &\quad \left. + \delta^{T+1} \left[\rho \left(\delta e^I V - \frac{c\,(e^I)^2}{2} \right) + (1-\rho) \left(\delta e^T V - \frac{c\,(e^T)^2}{2} \right) \right] + \frac{\delta^{T+2}}{1-\delta} \left(\delta e^T V - \frac{c\,(e^T)^2}{2} \right) \right]. \end{aligned}$$

This is smaller than U_S^R, therefore a renegotiation can never be optimal. ■

Before going on, note that a termination (threat) is required to prevent renegotiation.

If a team tried to implement such a scheme without a termination (i. e., any attempt for renegotiation causes a change of continuation play as laid out above), this could (and would) be subject to renegotiation as well.

In a next step, we show that such an equilibrium can also be made renegotiation-proof with respect to off-path deviations (and be weakly or strictly renegotiation-proof in the sense of Farrell and Maskin (1989)), implying that no subgame Pareto-dominates another.[18] In this case, free riding is not punished by a complete reversion to individual production. Instead, the relationship enters a punishment phase. In the punishment phase, both agents work on individual projects for a number of periods before resuming teamwork, where however the free rider is only allowed to exert an effort level below e^I. Naturally, this also must be incentive compatible for the free rider who always has the option to choose e^I when working on an individual project. Incentive compatibility for the free rider is achieved by a restart of the punishment phase, which implies an additional delay before teamwork is resumed. The number of periods with individual production is constructed such that the other, non-deviating, agent is indifferent compared to always exerting e^T in the future. Because we do not allow for transfers, Pareto-optimality therefore does not imply that no surplus is destroyed off the equilibrium path. The equilibrium is constructed such that the non-deviating agent is not worse off. The free-rider, on the other hand, is punished. Finally, note that it is of no relevance whether the punishment phase following free-riding is carried out in the current relationship, or if it leads to a termination with a potential new relationship starting with the punishment phase (which still requires sufficient transparency such that new partners are aware of a free-rider's past behavior).

Proposition 5 *For δ sufficiently high, there exists an equilibrium where e^{FB} is implemented in every period and which is renegotiation-proof in a sense that no subgame Pareto-dominates another subgame.*

Proof of Proposition 5. We construct an equilibrium where $e^{\mathrm{T}} = e^{\mathrm{FB}}$ is exerted in every period (besides the first period of the game where e^{I} is possible) on the

[18] According to Farrell and Maskin (1989), an equilibrium σ is weakly renegotiation-proof in the sense of Farrell and Maskin if no continuation equilibria σ^1 and σ^2 exist such that σ^1 strictly Pareto-dominates σ^2. It is strongly renegotiation-proof if none of its continuation equilibria is strictly Pareto-dominated by another weakly renegotiation-proof equilibrium.

equilibrium path. After one agent has deviated by not contributing team effort e^{T}, the game enters a punishment phase.[19] Because of renegotiation proofness, the non-deviating agent is not worse off at the beginning of the punishment phase than on the equilibrium path. On-path renegotiations are excluded by the strategies derived in the proof of Proposition 4, hence we assume that the conditions needed there are satisfied. This implies that we can also assume that a deviation by *both* players triggers a termination, and play in new matches is given by the strategies derived above. In the following, we therefore focus on unilateral deviations which involve free-riding on the other's effort.

It remains to show that given these requirements, there exist values of δ such that exerting effort $e^{\mathrm{T}} = e^{\mathrm{FB}}$ in every period is optimal. We start with constructing the punishment phase which is started after exactly one agent exerted less than e^{T} for the team project.

In the punishment phase, the non-deviating agents works solely on his own project for $T > 1$ periods, exerting effort e^{I}. In period $T+1$, he returns to teamwork with probability $1-\rho$ and works for another period on his own with probability ρ. From period $T+2$ onwards, both agents return to teamwork forever after. T and ρ are chosen such that the non-deviating agent's utility in the punishment phase is the same as on the equilibrium path. The off-path utility of the non-deviating party is

$$\beta\delta e^{\mathrm{I}}V - \frac{c\left(e^{\mathrm{I}}\right)^2}{2} + \delta\beta\left\{\left(\delta e^{\mathrm{I}}V - \frac{c\left(e^{\mathrm{I}}\right)^2}{2}\right)\frac{1-\delta^T}{1-\delta}\right.$$
$$\left. + \delta^T\left[\rho\left(\delta e^{\mathrm{I}}V - \frac{c\left(e^{\mathrm{I}}\right)^2}{2}\right) + (1-\rho)\left(\delta e^{\mathrm{T}}V - \frac{c\left(e^{\mathrm{T}}\right)^2}{2}\right)\right] + \frac{\delta^{T+1}}{1-\delta}\left(\delta e^{\mathrm{T}}V - \frac{c\left(e^{\mathrm{T}}\right)^2}{2}\right)\right\}.$$

For first-best effort, this becomes

$$\frac{\beta^2\delta^2V^2}{2c} + \delta\beta\frac{\delta^2V^2}{2c}\left[\frac{(2-\beta)\beta + \delta^T(1-\beta)^2 - \delta^T\rho(1-\beta)^2(1-\delta)}{1-\delta}\right].$$

On-path utility with $e^{\mathrm{I}} = e^{\mathrm{FB}} = \delta V/c$ equals

$$U^{\mathrm{T}} = \frac{\delta^2V^2}{2c}\left[\frac{-1+2\beta+\delta(1-\beta)}{1-\delta}\right].$$

[19] Without loss of generality, we assume that punishment after a unilateral deviation happens within a relationship and does not trigger a termination.

Therefore, ρ and T are characterized by

$$-1+\delta\left(1+\beta\right)=\beta\delta^{T+1}\left[1-\rho\left(1-\delta\right)\right]. \tag{19}$$

The deviating agent is only allowed to exert a reduced amount of effort on an individual project during the punishment phase. This amount is denoted $e^{\mathrm{P}}=\alpha e^{\mathrm{I}}$, with $\alpha\leq 1$. If the deviating agent does not follow through and chooses e^{I} instead of e^{P} during a punishment phase, the punishment phase starts over again.

Now the following conditions must hold: (i) It must be optimal to select e^{T} in every period on the equilibrium path. (ii) It must be optimal for a deviating agent to select e^{P} in every period of the punishment phase, and (iii) utility in the punishment phase must not be smaller than U^{I} – because an agent always has the option to go for individual production.

Before we derive the respective condition, we characterize the deviating agent's payoff in all relevant situations. The utility of the deviating agent at the beginning of the punishment phase equals

$$\begin{aligned}U^{\mathrm{P}}=\beta\delta e^{\mathrm{P}}V-\frac{c(e^{\mathrm{P}})^2}{2}+\delta\beta\Bigg\{&\left(\delta e^{\mathrm{P}}V-\frac{c\left(e^{\mathrm{P}}\right)^2}{2}\right)\frac{1-\delta^{T}}{1-\delta}\\&+\delta^{T}\left[\rho\left(\delta e^{\mathrm{P}}V-\frac{c\left(e^{\mathrm{P}}\right)^2}{2}\right)+\left(1-\rho\right)\left(\delta e^{\mathrm{T}}V-\frac{c\left(e^{\mathrm{T}}\right)^2}{2}\right)\right]+\frac{\delta^{T+1}}{1-\delta}\left(\delta e^{\mathrm{T}}V-\frac{c\left(e^{\mathrm{T}}\right)^2}{2}\right)\Bigg\}.\end{aligned}$$

Using (19), this becomes

$$U^{\mathrm{P}}=\beta\delta e^{\mathrm{P}}V-\frac{c(e^{\mathrm{P}})^2}{2}+\left(\delta e^{\mathrm{P}}V-\frac{c\left(e^{\mathrm{P}}\right)^2}{2}\right)+\frac{1}{1-\delta}\left(\delta e^{\mathrm{T}}V-\frac{c\left(e^{\mathrm{T}}\right)^2}{2}\right)\left[-1+\delta\left(1+\beta\right)\right].$$

Plugging in $e^{\mathrm{T}}=e^{\mathrm{FB}}=\delta V/c$, we get

$$U^{P}=\alpha\frac{\beta\delta^2V^2}{c}\left(1+\beta\left(1-\alpha\right)\right)+\frac{1}{1-\delta}\frac{\delta^2V^2}{2c}\left[-1+\delta\left(1+\beta\right)\right].$$

Deviating in the punishment phase (that is, selecting e^{I} for one period, followed by a re-start of the punishment phase), yields a utility of

$$\begin{aligned}U^{\mathrm{PD}}=\beta\delta e^{\mathrm{I}}V-\frac{c(e^{\mathrm{I}})^2}{2}+\delta\beta\Bigg\{&\left(\delta e^{\mathrm{P}}V-\frac{c\left(e^{\mathrm{P}}\right)^2}{2}\right)\frac{1-\delta^{T+1}}{1-\delta}\\&+\delta^{T+1}\left[\rho\left(\delta e^{\mathrm{P}}V-\frac{c\left(e^{\mathrm{P}}\right)^2}{2}\right)+\left(1-\rho\right)\left(\delta e^{\mathrm{T}}V-\frac{c\left(e^{\mathrm{T}}\right)^2}{2}\right)\right]+\frac{\delta^{T+2}}{1-\delta}\left(\delta e^{\mathrm{T}}V-\frac{c\left(e^{\mathrm{T}}\right)^2}{2}\right)\Bigg\}.\end{aligned}$$

Using (19), this becomes

$$U^{\mathrm{PD}} = \beta\delta e^{\mathrm{I}} V - \frac{c(e^{\mathrm{I}})^2}{2} + \frac{\delta}{1-\delta}\left\{\left(\delta e^{\mathrm{P}} V - \frac{c\,(e^{\mathrm{P}})^2}{2}\right)(1+\beta)(1-\delta) + \left(\delta e^{\mathrm{T}} V - \frac{c\,(e^{T})^2}{2}\right)(-1+\delta\,(1+\beta))\right\}.$$

Plugging in $e^{\mathrm{T}} = e^{\mathrm{FB}} = \delta V/c$, we get

$$U^{PD} = \frac{\beta^2\delta^2 V^2}{2c} + \delta\alpha\frac{\beta\delta^2 V^2}{2c}(2-\alpha\beta)(1+\beta) + \frac{\delta^2 V^2}{2c}\frac{\delta}{1-\delta}(-1+\delta\,(1+\beta)).$$

The utility of free-riding amounts to

$$\begin{aligned} U^{\mathrm{D}} &= \beta\delta V\left(\frac{e^{\mathrm{T}}}{2} + e^{\mathrm{I}}\right) - \frac{c\,(e^{\mathrm{I}})^2}{2} + \beta\delta\left\{\left(\delta e^{\mathrm{P}} V - \frac{c\,(e^{\mathrm{P}})^2}{2}\right)\frac{1-\delta^{T+1}}{1-\delta}\right. \\ &\quad \left. + \delta^{T+1}\left[\rho\left(\delta e^{\mathrm{P}} V - \frac{c\,(e^{\mathrm{P}})^2}{2}\right) + (1-\rho)\left(\delta e^{\mathrm{T}} V - \frac{c\,(e^{\mathrm{T}})^2}{2}\right)\right] + \frac{\delta^{T+2}}{1-\delta}\left(\delta e^{\mathrm{T}} V - \frac{c\,(e^{\mathrm{T}})^2}{2}\right)\right\} \\ &= \beta\delta V\frac{e^{\mathrm{T}}}{2} + U^{\mathrm{PD}}. \end{aligned}$$

Finally,

$$U^{\mathrm{I}} = \frac{\beta^2\delta^2 V^2}{2c}\frac{1+\delta\,(1-\beta)}{1-\delta} \quad \text{and}$$
$$U^{\mathrm{T}} = \frac{\delta^2 V^2}{2c}\left[\frac{-1+2\beta+\delta\,(1-\beta)}{1-\delta}\right].$$

Now, we can derive three conditions that make sure e^{FB} can be implemented in every period. The on-path (IC) constraint requires $U^{\mathrm{T}} \geq U^{\mathrm{D}}$, which is equal to

$$[-1+\delta\,(1+\beta)] \leq \frac{\beta-(1-\delta)\,(1+\beta^2)-\delta\alpha\beta\,(2-\alpha\beta)\,\beta}{\delta\,(1-\alpha\beta)^2}. \tag{onIC}$$

The off-path (IC) constraint requires $U^{\mathrm{P}} \geq U^{\mathrm{PD}}$, which is equal to

$$[-1+\delta\,(1+\beta)] \geq \frac{\beta^2\,(1-\alpha)^2}{(1-\alpha\beta)^2}. \tag{offIC}$$

Finally, $U^{\mathrm{P}} \geq U^{\mathrm{I}}$ and $U^{\mathrm{PD}} \geq U^{\mathrm{I}}$ must hold. Because of (offIC), though, only the

second condition is relevant, which becomes

$$-1-\beta^2(2-\beta)+\left[\alpha\beta(2-\alpha\beta)+\delta(1-\alpha\beta)^2\right](1+\beta)\geq 0.$$

It is our objective to find one example for such an equilibrium. Therefore, we will show that all these conditions hold for $\alpha=0$ if δ is sufficiently large.

With $\alpha=0$, (onIC) becomes $-1\leq(\beta-(1-\delta)(1+\beta^2))/\delta-\delta(1+\beta)$. The first-order derivative of the right-hand side with respect to δ equals $(1-\beta+\beta^2)/\delta^2-(1+\beta)$, the second-order derivative $-2\frac{1-\beta+\beta^2}{\delta^3}$. Therefore, the right-hand side of (onIC) is maximized by $\delta^*=\sqrt{\frac{1-\beta+\beta^2}{(1+\beta)}}<1$. Since (onIC) becomes $1+\beta\leq 1+\beta$ for $\delta=1$, it also holds for δ^*. All this implies that there exists a $\tilde{\delta}<\delta^*$ such that (onIC) is satisfied for $\delta\in[\tilde{\delta},1]$.

With $\alpha=0$, equation (offIC) becomes $[-1+\delta(1+\beta)]\geq\beta^2$. This holds for $\delta\geq(1+\beta^2)/(1+\beta)$, with $(1+\beta^2)/(1+\beta)<1$. Finally, $U^{\mathrm{PD}}\geq U^{\mathrm{I}}$ becomes $-1-\beta^2(2-\beta)+\delta(1+\beta)\geq 0$. This holds for $\delta(1+\beta)\geq\frac{1+\beta^2(2-\beta)}{(1+\beta)}$, with $\frac{1+\beta^2(2-\beta)}{(1+\beta)}<1$. Hence, we have shown that for δ sufficiently large, there exists a renegotiation-proof equilibrium with $\alpha=0$ where $e^{\mathrm{T}}=e^{\mathrm{FB}}$ can be implemented in every period. ■

Therefore, under the conditions derived above, we are able to construct an equilibrium that is robust to allowing for on- as well as off-path renegotiation.

B.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,

$$\left(\delta e_1^{\mathrm{T}} \frac{V}{2} - \frac{c\,(e_1^{\mathrm{T}})^2}{2}\right) - \frac{\delta^2 V^2}{2c} + \frac{\delta}{1-\delta}\left[\left(\delta\left(e_1^{\mathrm{T}} + e_2^{\mathrm{T}}\right)\frac{V}{2} - \frac{c\,(e_1^{\mathrm{T}})^2}{2}\right) - \frac{\delta^2 V^2}{2c}\right] \geq 0. \tag{20}$$

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

$$e_2^{\mathrm{T}} \geq \frac{V}{c} \tag{21}$$

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

$$-1 + \delta\left(1 + \beta_2\delta\left(\delta - \beta_2 - \beta_2\delta + \beta_2^2\delta\right)\right) \geq 0. \tag{22}$$

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.

B.3 Imperfect Public Monitoring

In this section, we sketch the implications of mutual monitoring being imperfect. 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 conditions 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,

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

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

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

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

B.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_tV 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

$$\left(\beta\delta e^{\mathrm{T}}\frac{V}{2} - C(e^{\mathrm{T}})\right) - \left(\beta\delta e^{\mathrm{I}}V - C(e^{\mathrm{I}})\right)$$
$$+\frac{\beta\delta}{1-\delta}\Big[\left(\delta e^{\mathrm{T}}V - C(e^{\mathrm{T}})\right) - \left(\delta e^{\mathrm{I}}V - C(e^{\mathrm{I}})\right)\Big] \geq 0. \qquad \text{(IC)}$$

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

Proof of Lemma 3. Note that in this case, $e^{\mathrm{I}} = e^{\mathrm{FB}}$. The constraint becomes

$$\left(\delta e^{\mathrm{T}}\frac{V}{2} - C(e^{\mathrm{T}})\right) - \left(\delta e^{\mathrm{FB}}V - C(e^{\mathrm{FB}})\right)$$
$$+\frac{\delta}{1-\delta}\Big[\left(\delta e^{\mathrm{T}}V - C(e^{\mathrm{T}})\right) - \left(\delta e^{\mathrm{FB}}V - C(e^{\mathrm{FB}})\right)\Big] \geq 0 \qquad (23)$$

Now, e^{FB} maximizes $\delta eV - C(e)$, hence $\left(\delta e^{\mathrm{T}}V - C(e^{\mathrm{T}})\right) - \left(\delta e^{\mathrm{FB}}V - C(e^{\mathrm{FB}})\right) \leq 0$ for any e^{T}; furthermore, $\left(\delta e^{\mathrm{T}}\frac{V}{2} - C(e^{\mathrm{T}})\right) - \left(\delta e^{\mathrm{FB}}V - C(e^{\mathrm{FB}})\right) < 0$ 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 4 *No effort level $e^T \leq e^I$ can be enforced within a team.*

Proof of Lemma 4. 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), $\left(\delta e^{\mathrm{T}}V - C(e^{\mathrm{T}})\right) - \left(\delta e^{\mathrm{I}}V - C(e^{\mathrm{I}})\right) \leq 0$ for $e^{\mathrm{T}} \leq e^{\mathrm{I}}$; because e^{I} maximizes $\beta\delta eV - C(e)$ the first line of the (IC) constraint, $\beta\delta e^{\mathrm{T}}V/2 - C(e^{\mathrm{T}}) - \left(\beta\delta e^{\mathrm{I}}V - C(e^{\mathrm{I}})\right)$, 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 6 *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 of Proposition 6. The second line of (IC), $\left(\delta e^{\mathrm{T}} V - C(e^{\mathrm{T}})\right) - \left(\delta e^{\mathrm{I}} V - C(e^{\mathrm{I}})\right)$, is strictly positive for any $\beta < 1$ and $e^{\mathrm{I}} < e^{\mathrm{T}} \leq e^{\mathrm{FB}}$. Following Lemmas 3 and 4, 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 7 *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 of Proposition 7. 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

$$\delta\left(e^{\mathrm{T}}\frac{V}{2} - e^{\mathrm{I}}V\right) - \beta\frac{\delta}{1-\delta}\frac{de^{\mathrm{I}}}{d\beta}\delta\left(1-\beta\right)V + \frac{\delta}{1-\delta}\left[\left(\delta e^{\mathrm{T}}V - C(e^{\mathrm{T}})\right) - \left(\delta e^{\mathrm{I}}V - C(e^{\mathrm{I}})\right)\right] \tag{24}$$

For all e^{T} with $0 < e^{\mathrm{T}} \leq e^{\mathrm{FB}}$, (24) 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 (24) 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

the (IC) constraint 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

$$\frac{de^{\mathrm{I}}}{d\beta}\frac{\delta}{1-\delta}V\left[3\beta\delta-1-\delta\right], \tag{25}$$

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 (25) 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 (24) 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 (24) equals zero and where (25) is non-positive. Finally, $\beta^{\max}$ is unique because of the non-existence of a minimum. ■

References

Abreu, D. (1988): "On the Theory of Infinitely Repeated Games with Discounting," *Econometrica*, 56(6), 1713–1733.

Akerlof, G. A. (1982): "Labor Contracts as Partial Gift Exchange," *Quarterly Journal of Economics*, 97(4), 543–569.

Alchian, A. A., and H. Demsetz (1972): "Production, Information Costs, and Economic Organization," *American Economic Review*, 62(5), 777–795.

Ashraf, N., D. Karlan, and W. Yin (2006): "Tying Odysseus to the Mast:

Evidence from a Commitment Savings Product in the Philippines," *Quarterly Journal of Economics*, 121(2), 635–672.

Augenblick, N., M. Niederle, and C. Sprenger (2013): "Working Over Time: Dynamic Inconsistency in Real Effort Tasks," Working Paper.

Bar-Isaac, H. (2007): "Something to Prove: Reputation in Teams," *RAND Journal of Economics*, 38(2), 495–511.

Baron, J., and D. M. Kreps (1999): *Strategic Human Resources: Frameworks for General Managers*. Wiley & Sons, Hoboken, NJ.

Bartling, B. (2011): "Relative Performance or Team Evaluation? Optimal Contracts for Other-Regarding Agents," *Journal of Economic Behavior and Organization*, 79(3), 183–193.

Basu, K. (2011): "Hyperbolic Discounting and the Sustainability of Rotational Savings Arrangements," *American Economic Journal: Microeconomics*, 3(4), 143–171.

Battaglini, M., R. Benabou, and J. Tirole (2005): "Self-Control in Peer Groups," *Journal of Economic Theory*, 123(2), 105–134.

Bauer, M., J. Chytilova, and J. Morduch (2012): "Behavioral Foundations of Microcredit: Experimental and Survey Evidence from Rural India," *American Economic Review*, pp. 1118–1139.

Bernheim, B. D., D. Ray, and S. Yeltekin (2013): "Poverty and Self-Control," Working Paper.

Bond, P., and G. Sigurdsson (2013): "Commitment Contracts," Working Paper.

Bull, C. (1987): "The Existence of Self-Enforcing Implicit Contracts," *Quarterly Journal of Economics*, 102(1), 147–159.

Che, Y.-K., and S.-W. Yoo (2001): "Optimal Incentives for Teams," *American Economic Review*, 91(3), 525–541.

Corts, K. S. (2007): "Teams versus Individual Accountability: Solving Multitask Problems through Job Design," *RAND Journal of Economics*, 38(2), 467–479.

DellaVigna, S., and U. Malmendier (2004): "Contract Design and Self-Control: Theory and Evidence," *Quarterly Journal of Economics*, 119(2), 353–402.

——— (2006): "Paying Not to Go to the Gym," *American Economic Review*, 96(3), 694–719.

Englmaier, F., and S. Leider (2012): "Contractual and Organizational Structure with Reciprocal Agents," *American Economic Journal: Microeconomics*, 4(2), 146–183.

Falk, A., and U. Fischbacher (2006): "A Theory of Reciprocity," *Games and Economic Behavior*, 54(2), 293–315.

Farrell, J., and E. Maskin (1989): "Renegotiation in Repeated Games," *Games and Economic Behavior*, 1, 327–360.

Forsyth, D. R. (2008): "Self-Serving Bias," .

Hestermann, N., and Y. Le Yaouanq (2016): "It's Not My Fault! Ego, Excuses and Perseverance," Working Paper.

Holmstrom, B. (1982): "Moral Hazard in Teams," *Bell Journal of Economics*, 13(2), 324–340.

Itoh, H. (1991): "Incentives to Help in Multi-Agent Situations," *Econometrica*, 59(3), 611–636.

Jones, E. E., and R. E. Nisbett (1987): "The Actor and the Observer: Divergent Perceptions of the Causes of Behavior," in *Attribution: Perceiving the causes of behavior*, ed. by E. E. Jones, D. E. Kanouse, H. H. Kelley, R. E. Nisbett, S. Valins, and B. Weiner, pp. 79 – 94. Lawrence Erlbaum associates.

Kandel, E., and E. P. Lazear (1992): "Peer Pressure and Partnerships," *Journal of Political Economy*, 100(4), 801–817.

Kaur, S., M. Kremer, and S. Mullainathan (2010): "Self -Control and the Development of Work Arrangements," *American Economic Review Papers and Proceedings*, 100(2), 624–628.

——— (2013): "Self-Control at Work," Working Paper.

Kim, J.-H., and N. Vikander (2013): "Team-Based Incentives in Problem-Solving Organizations," *Journal of Law, Economics, and Organizations*, forthcoming.

Krusell, P., and J. Anthony A. Smith (2003): "Consumption-Savings Decisions with Quasi-Geometric Discounting," *Econometrica*, 71(1), 365–375.

Kvaløy, O., and T. E. Olsen (2006): "Team Incentives in Relational Employment Contracts," *Journal of Labor Economics*, 24(1), 139–169.

Laibson, D. (1994): "Self-Control and Saving," Working Paper.

——— (1997): "Golden Eggs and Hyperbolic Discounting," *Quarterly Journal of Economics*, 112(2), 443–477.

Landeo, C. M., and K. E. Spier (2015): "Incentive contracts for teams: Experimental evidence," *Journal of Economic Behavior & Organization*, 119, 496 – 511.

Lau, R. R., and D. Russell (1980): "Attribution in the Sports Pages," *Journal of Personality and Social Psychology*, 39(1), 29–38.

Lazear, E. P., and K. L. Shaw (2007): "Personnel Economics: The Economist's View of Human Resources," *Journal of Economic Perspectives*, 21(4), 91–114.

Levin, J. (2003): "Relational Incentive Contracts," *American Economic Review*, 93(3), 835–857.

MacLeod, W. B., and J. M. Malcomson (1989): "Implicit Contracts, Incentive Compatibility, and Involuntary Unemployment," *Econometrica*, 57(2), 447–480.

Mohnen, A., K. Pokorny, and D. Sliwka (2008): "Transparency, Inequity Aversion, and the Dynamics of Peer Pressure in Teams: Theory and Evidence," *Journal of Labor Economics*, 26(4), 693–720.

Moskowitz, G. B. (2005): *Social Cognition: Understanding Self and Others*. Guilford Press.

Mukherjee, A., and L. Vasconcelos (2011): "Optimal Job Design in the Presence of Implicit Contracts," *RAND Journal of Economics*, 42(1), 44–69.

O'Donoghue, T., and M. Rabin (1999): "Doing it Now or Later," *American Economic Review*, 89(1), 103–124.

——— (2001): "Choice and Procrastination," *Quarterly Journal of Economics*, 116(1), 121–160.

Phelps, E., and R. Pollak (1968): "On Second-best National Saving and Game-Equilibrium Growth," *Review of Economic Studies*, 35(2), 185–199.

Rayo, L. (2007): "Relational Incentives and Moral Hazard in Teams," *Review of Economic Studies*, 74(3), 937–963.

Stewart, A. E. (2005): "Attributions of responsibility for motor vehicle crashes," *Accident Analysis & Prevention*, 37(4), 681 – 688.

Strotz, R. (1955): "Myopia and Inconsistency in Dynamic Utility Maximization," *Review of Economic Studies*, 23(3), 165–180.

Zuckerman, M. (1979): "Attribution of Success and Failure Revisited, or: The Motivational Bias is Alive and Well in Attribution Theory," *Journal of Personality*, 47(2), 245–287.