Greiff, Matthias

Working Paper

Integrating affective responses into game theory: A dual selves model


Provided in Cooperation with:
Faculty of Business Administration and Economics, University of Marburg


This Version is available at:
http://hdl.handle.net/10419/119454

Terms of use:
Documents in EconStor may be saved and copied for your personal and scholarly purposes.
You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.
If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

www.econstor.eu
Integrating Affective Responses into Game Theory: A Dual Selves Model

This paper can be downloaded from http://www.uni-marburg.de/fb02/makro/forschung/magkspapers/index_html%28magks%29

Coordination: Bernd Hayo • Philipps-University Marburg
School of Business and Economics • Universitätsstraße 24, D-35032 Marburg
Tel: +49-6421-2823091, Fax: +49-6421-2823088, e-mail: hayo@wiwi.uni-marburg.de
Integrating Affective Responses into Game Theory: 
A Dual Selves Model

Matthias Greiff

June 9, 2015

Abstract

This paper develops a method to integrate affective responses into game theoretical models. We illustrate our method in a team production framework. The model analyzes how concave and convex status preferences for esteem solve the problem of team production under complete and incomplete information about workers’ abilities. Using a dual selves model, we model the choice of effort as a deliberative decision and the expression of esteem as an affective response. Modeling an individual’s affective system as a separate player allows us to apply standard game-theoretic solution concepts to analyze affective responses.

JEL Codes: D03, D21, D8, M5
1 Rational Decisions and Affective Responses

Most of economics is based on the assumption that behavior can be modeled as the result of the maximization of a well-defined utility function. One way to incorporate emotions into economic analysis is to include the hedonic impact of emotions in the utility function. Kräkel (2008), for example, presents a theoretical model in which emotion is an additional argument in the utility function and argues that these emotions are triggered if team members compare their performance with the performance of co-workers. This approach, however, focuses on how emotions affect behavior but remains silent about the source of emotions.

Empirical evidence for the role of emotions comes from ultimatum and power-to-take games, where it is often argued that negative emotions determine punishment behavior. Based on experimental and neuroeconomic evidence, Pillutla and Murnighan (1996) and Sanfey et al. (2003) argue that the rejection of low offers in the ultimatum game is driven by negative emotions which are triggered by perceived unfairness.

In the power-to-take game (Bosman and van Winden, 2002), two players, proposer and responder, have endowments $e_p$ and $e_r$. The proposer moves first by choosing the take rate, $t$. After observing the take rate, the responder chooses a destruction rate, $d$. The destruction rate determines the fraction of $e_r$ that gets destroyed before the fraction $t$ of the responder’s remaining endowment is transferred to the proposer. Payoffs are $e_p + t(1 - d)e_r$ for the proposer and $(1 - d)(1 - t)e_r$. While in the ultimatum game punishment is a binary decision (reject or accept), punishment in the power-to-take game is more nuanced since $d$ can take any value between 0 and 1. Bosman and van Winden (2002) argue that the choice of the destruction rate is driven by the intensity of negative emotions.

The results indicate how emotions might affect behavior. In both games, emotions are expressed by responders’ decisions. Do these decisions result from a rational cost-benefit calculus or are they triggered by the affective system? In the power-to-take game, decision times for responders who chose $d = 1$ are significantly lower than the decision times for responders who chose to destroy a positive amount but less than everything. This can be interpreted as evidence for a fast and affective system which, together with a slow and deliberative system, determines behavior. If emotional arousal is low, decision times are high because of the interaction between the rational system’s prediction (destroy nothing) and the affective system’s prediction (destroy something), but if emotional arousal is strong, decision times are low because the affective system takes over (van Winden, 2007). Based on these results, van Winden (2007) claims that emotion-promoted decisions should not be conceptualized as resulting from a rational cost-benefit calculus but as affective responses. This is similar to Frank (1988), who argues that emotions cannot be

---

1In the ultimatum game there are two players, proposer and responder. The proposer has endowment $e$ and the responder has no endowment. The proposer offers a share $s$ of her endowment to the responder. If the responder accepts, the payoffs are $(1 - s)e$ and $se$; otherwise both players receive nothing.
simulated, and to Brennan and Pettit (2000), who argue that “esteem has to be generated involun-
tarily by how things seem to the estimator”, and that “[e]steem cannot be provided or passed on
by way of voluntary choice” (p. 89). Put bluntly, emotions like esteem are special goods whose
production and supply is beyond the rational actor model.

If the expression of emotions is beyond the rational actor model, it is not clear how it can be
integrated into game-theoretic models. In this paper, we focus on one particular emotion, esteem,
and show how emotionally-prompted decisions, like the expression of esteem, can be integrated
into game-theoretic analysis. To illustrate our approach, we use a simple model of team produc-
tion. In line with existing research, we assume that esteem is an argument in the team members’
utility functions. The expression of esteem, however, is not a decision resulting from rational
utility-maximizing behavior but an affective response. To model affective responses, we choose a
dual selves perspective in which each individual consists of a rational and an emotional self. In a
stylized way, the emotional self takes the affective decisions without strategic considerations, and
the rational self makes the strategic decisions taking emotions into account. Extending the set of
players allows us (i) to study the expression of esteem using standard game-theoretical methods,
(ii) to investigate how the expression of emotions and its effects depend on the information being
available, and (iii) to analyze which preferences are necessary for esteem to be effective in solving
the problem of team production.

2 Team Production and Esteem Incentives

How to motivate team members? To study this question we use a public goods framework in
which endowments are interpreted as ability and contributions are interpreted as effort (Alchian
and Demsetz, 1972). Team members’ effort levels generate a public good which affects team mem-
bers’ utilities. Within this framework, we show how the problem of team production can be solved
if workers have preferences for esteem and react to each others’ relative contributions by express-
ing esteem. The expression is not a rational decision but an affective response.

The role of esteem has been investigated in principal agent settings (Ellingsen and Johann-
esson, 2007, 2008) and in public good settings (Hollander, 1990; Kandel and Lazear, 1992; Brenn-
an and Pettit, 2000; Cowen, 2002; Brennan and Brooks, 2007). In order to focus on behavior within
the team, we abstract from agency problems and do not consider interactions between team and
employer. Instead, we focus on team production and investigate the role of esteem incentives

---

2 There is a large literature on dual process models, surveyed in O’Donoghue and Loewenstein (2007), Evans (2008)
and Alós-Ferrer and Strack (2014). In dual process models, behavior is determined by the interaction between delibera-
tive and affective processes. We model each process as a separate player.

3 Related to non-monetary incentives and agency problems is, of course, the literature on contractual incompleteness,
reciprocity and gift-exchange in the workplace. This literature finds that the employer’s perceived kindness matters,
and that perceived kindness depends not only on the wage but also on intentions (Charness, 2004), or the nature of the
gift (Kube et al., 2012).
in teams with heterogeneous workers. Within a simple static model with two workers, we analyze how esteem incentives solve the problem of team production under complete and incomplete information about workers’ abilities.

The idea that esteem incentives increase prosocial behavior is not new. Evidence from the lab and from the field suggests that non-monetary incentives, like esteem, social approval, respect, or peer evaluations, increase prosocial behavior and social capital (Gächter and Fehr, 1999; Masclet et al., 2003; Kosfeld and Neckermann, 2010; Neckermann and Frey, 2013; Greiff and Paetzel, 2015). A theory of esteem has been developed by Brennan and Pettit (Pettit, 1990; Brennan and Pettit, 1993, 2004) and has been applied to public goods provision in Cowen (2002) and Brennan and Brooks (2007). The effect of esteem incentives on behavior might be affected by social comparison, for example, if the motivational strength of esteem depends not only on the amount of esteem a worker receives but on her relative standing among peers. Such a preference for status seeking has been documented theoretically and empirically (Frank, 1985; Loch et al., 2000; Zizzo, 2002; Huberman et al., 2004; Rege, 2008). Closely related is the literature on information provision and performance. Using data from laboratory experiments, Falk and Ichino (2006) and Kuhnen and Tymula (2012) show that a participant’s individual performance increases if she is informed about other participants’ performances. The effect of information about the performance of peers on individual performance was confirmed in the field, where Azmat and Iriberri (2010) and Tran and Zeckhauser (2012) showed that relative performance feedback induces students to perform better. Hence, the evidence suggests that information about individual performance relative to others’ performance matters, even if payment is unrelated to performance. Taking it as given that workers have a competitive preference for esteem, we analyze the interplay between esteem incentives and prosocial behavior. Specifically, our goal is to explain how a competitive preference for esteem interacts with information about workers’ heterogeneity.

In a nutshell, the model is as follows: Workers have heterogeneous abilities. Information about abilities is either public or private information. Effort is observable by team members, but not contractible. After contributing to team production, esteem is expressed by the workers’ emotional systems. When workers’ abilities are common knowledge, esteem can solve the problem of team production. With incomplete information, however, esteem solves the problem of team production only if two conditions are satisfied: First, a worker expressing esteem compares a coworker’s relative effort to the relative effort of a reference group. We call this the social comparison condition. Second, utility increases with the difference of esteem one receives and esteem received by others. We call this the status condition.

Closely related to our model are Hollander (1990), Kandel and Lazear (1992), Cowen (2002) and Brennan and Brooks (2007). Kandel and Lazear analyze how peer pressure affects voluntary contributions to a public good. In their model, exerting peer pressure is a conscious decision, while in Hollander (1990), Cowen (2002) and Brennan and Brooks (2007) the supply of esteem is
an affective response. In this respect, our model is closest to Hollander’s model. Although we model the expression of esteem explicitly, the expression of esteem is not the result of a conscious decision resulting from utility maximization but an affective response triggered automatically by the emotional system which we model as an additional player. Hollander (1990), Kandel and Lazear (1992), Cowen (2002) and Brennan and Brooks (2007) focus on the relation between esteem and voluntary contributions to a public good, but since they do not allow for heterogeneous players and incomplete information, they are unable to explore the interaction of esteem-based contributions and incomplete information.

3 Theoretical Model

We consider a two-player game which consists of three stages. In the first stage, nature decides about workers’ types. Workers’ types are determined by independent draws from the lottery \( \mathcal{L} := (1 - p) \cdot \theta^l + p \cdot \theta^h \) with \( 0 < \theta^l < \theta^h \) and \( p > 0 \). A worker’s type is her ability. Let \( \hat{\theta} := (1 - p) \theta^l + p \theta^h \) denote expected ability. Hence, there are two workers, \( i \) and \( j \), with abilities \( \theta_i \) and \( \theta_j \). The probability \( p \) and the values for \( \theta^l \) and \( \theta^h \) are common knowledge. Abilities can be common knowledge (section 5) or private information (section 6).

In the second stage, all workers simultaneously choose effort, \( e_i (i \in \{1, 2\}) \). The choice of effort is a rational decision. We assume that efforts are observable but not verifiable. For simplicity, we assume that effort can be either low or high, \( e_i \in \{e^l, e^h\} \) for \( i \in \{1, 2\} \). This implies that a worker with low ability will always choose low effort, and that only workers with high ability can choose whether to exert low or high effort. For worker \( i \), the monetary payoff from the output of team production is given by \( \pi_i(e_1, e_2) \) for \( i \in \{1, 2\} \). Let the cost of effort be given by \( \Phi(e_i/\theta_i) \) and assume that the cost of effort is strictly increasing in relative effort, \( e_i/\theta_i \).

In the third stage, workers are informed about chosen efforts and each worker’s emotional system reacts to this information by publicly expressing esteem, which is modeled as an affective response.

3.1 Preferences (without esteem)

Consider the case without esteem where workers have preferences over monetary income and cost of effort. In order to simplify the problem, we assume that these preferences are separable and represented by \( u_i = \pi_i(e_1, e_2) - \Phi(e_i/\theta_i) \) where \( \pi_i(e_1, e_2) = e_1 + e_2 \) is the utility from the output of team production and \( \Phi(e_i/\theta_i) \) is the disutility of effort.\(^4\) Disutility of effort is type-dependent with

\[
\Phi \left( \frac{e_i}{\theta_i} \right) = \begin{cases} 
\gamma & \text{if } \frac{e_i}{\theta_i} = 1 \\
0 & \text{else}.
\end{cases}
\]

\(^4\)Here we depart from Alchian and Demsetz (1972), who assume non-separability.
Subgame \((i, j)\) Probability
---
1 \((\theta_i^l, \theta_j^l)\) \((1 - p)^2\)
2 \((\theta_i^l, \theta_j^h)\) \((1 - p)p\)
3 \((\theta_i^h, \theta_j^l)\) \(p(1 - p)\)
4 \((\theta_i^h, \theta_j^h)\) \(p^2\)

Table 1: Subgames in the game with complete information and probabilities for reaching them.

We assume \(0 < e^h - e^l \leq c \leq 2(e^h - e^l)\) so that each worker has an incentive to shirk but welfare is maximized when both workers choose not to shirk.

### 3.2 A Reference Model without Esteem

First we look at a simplified game where stage three – the affective response – is absent. Figures 1 and 2 show stages one and two in extensive form. Nature decides about the distribution of abilities before workers decide about effort. Table 1 summarizes the four subgames, the corresponding abilities and the probabilities for reaching each subgame. In the games in Figures 1 and 2 nodes \(E_1 - E_9\) are endnodes and the numbers at these nodes denote utilities derived from the monetary payoffs resulting from team production and costs of effort.

Figure 1 depicts the game in which abilities are common knowledge. A strategy consists of a vector \((X, Y)\). \(X\) denotes a worker’s effort when her ability is high and the other worker’s ability is low. \(Y\) denotes a worker’s effort when her own and the other worker’s ability is high. Let \(S_1\) be the set of all four pure strategies in the game with complete information. For abilities \((\theta^h, \theta^h)\) and without esteem, this is a prisoner’s dilemma, as we will see later.

Figure 2 depicts the game in which abilities are private information. A worker cannot condition her effort on the other worker’s ability. A strategy is denoted by \(Z\) and denotes a worker’s effort when her own ability is high. Let \(S_2\) be the set of all two pure strategies in the game with incomplete information.

In this reference model without esteem there is no “market” for esteem. Esteem is not supplied because no worker expresses esteem, and esteem is not in demand because workers have no preference for esteem. Hence, with public and private information about abilities, always choosing low effort is the dominant strategy in this reference model.
Figure 1: Game without esteem when abilities are public information.

Figure 2: Game without esteem when abilities are private information.
4 Incorporating Affective Responses

In this section we extend the games depicted in Figures 1 and 2 to allow for the expression of esteem. We add a third stage in which esteem is expressed through affective responses. The third stage of the extended game starts at nodes $E1 - E9$, i.e., these nodes are not endnodes (as in section 3.2) but nodes at which esteem is expressed.

4.1 Expressing Esteem

To incorporate affective responses into our model, we extend the set of players. The expression of esteem can be thought of as a behavioral response triggered by a player’s emotional system. Assume that there are two additional players, $A_1$ and $A_2$. Player $A_1$ is worker 1’s emotional system and expresses esteem $s_1$. Player $A_2$ is worker 2’s emotional system and expresses esteem $s_2$.

The expression of esteem happens in stage three, after both workers choose their effort levels. Players $A_1$ and $A_2$ simultaneously express esteem. We assume that esteem is expressed costlessly and that the strength of the affective response depends not only on a worker’s relative effort but also on the relative effort of some reference group (Brennan and Pettit, 2000, 80-81). Since we consider a team of two heterogeneous workers, we take the other worker as reference group. The stimulus power prompted by the choices of effort depends on both workers’ relative efforts. Assuming linearity, player $A_i$’s affective response is given by:

$$s_i \left( \frac{e_i}{\theta_i}, \frac{e_j}{\theta_j} \right) = \beta_1 \frac{e_j}{E_i[\theta_j|e_j]} - \beta_2 \frac{e_i}{\theta_i} \text{ for } i, j \in \{1, 2\}, \ i \neq j,$$

where $\beta_1$ and $\beta_2$ are nonnegative parameters. $E_i[\theta_j|e_j]$ is $i$’s expectation about $j$’s ability conditional on $j$’s effort. If abilities are common knowledge, $E_i[\theta_j|e_j] = \theta_j$ (for $i, j \in \{1, 2\}, i \neq j$). Depending on the parameters $\beta_1$ and $\beta_2$, we can distinguish three different affective responses.

(i) $\beta_1 > 0$ and $\beta_2 = 0$, (“Pure Respect”). This is the case if $A_i$ expresses respect for the other’s relative effort. The affective response is nonnegative and proportional to the other worker’s relative effort, $s_i = \beta_1(e_j/\theta_j)$. However, it is independent of the effort level chosen by the worker whose affective response we consider ($e_i/\theta_i$), so the social comparison condition is not fulfilled. Pure respect is an incentive that is supplied by one worker’s emotional system for another worker (e.g., supplied by player $A_1$ for worker 2).

(ii) $\beta_1 = 0$ and $\beta_2 > 0$, (“Pure Pride”). The affective response is nonpositive and inversely proportional to the effort level chosen by the worker whose affective response we consider. It is independent of the other’s effort, $s_i = -\beta_2(e_i/\theta_i)$. We can interpret $-s_i$ as $i$’s pride. While pure respect is expressed for another worker’s achievement, a worker derives pure pride from her own achievement. Pure pride is an incentive that is supplied by one worker’s emotional system for the worker herself (e.g., supplied by player $A_1$ for worker 1).
(iii) $\beta_1 > 0$ and $\beta_2 > 0$, ("Respect and Pride"). Player $A_i$’s affective response will be stronger if $j$’s relative effort is higher, or if $i$’s relative effort is lower (implying that the difference between $j$’s and $i$’s relative effort is higher). The parameter $\beta_2$ indicates the importance of the comparison to the reference group.

We assume that $\frac{\partial u_i}{\partial (c_i/\theta_i)} = \beta_1 + \alpha \beta_2 > 0$ so that a higher relative effort causes an increase in utility from esteem.\(^5\)

4.2 Preferences of $A_1$ and $A_2$

In order to derive $A_1$’s and $A_2$’s decisions from utility-maximization, we assume that $A_1$ and $A_2$ are expected utility maximizers for whom utilities depend on the accuracy of the information they use to express esteem, but not on the affective response per se. More formally, player $A_1$ has only one move which consists of choosing the probability $p_{1,2} \in [0, 1]$ with which worker 2 has high ability. This probability is then used to calculate the conditional expectation $E_1[\theta_2|e_2]$ which enters $A_1$’s affective response (equation 1). Similarly for player $A_2$. Affective responses affect both workers’ utilities, which are derived in the next section, but not $A_1$’s and $A_2$’s utilities, which are given by

$$
\begin{align*}
    u_{A_1}(p_{1,2}, \theta_2) &= -(I_2 - p_{1,2})^2 - [(1 - I_2) - (1 - p_{1,2})]^2, \quad \text{and} \\
    u_{A_2}(p_{2,1}, \theta_1) &= -(I_1 - p_{2,1})^2 - [(1 - I_1) - (1 - p_{2,1})]^2.
\end{align*}
$$

Here, $p_{1,2}$ is $A_1$’s belief that worker 2 has high ability, and $I_2$ is an indicator variable with $I_2 = 1$ only if worker 2 has high ability, $\theta_2 = \theta^h$. Similarly, $p_{2,1}$ is $A_2$’s belief that worker 1 has high ability, and $I_1$ is an indicator variable with $I_1 = 1$ only if worker 1 has high ability, $\theta_1 = \theta^h$.

With endowments being common knowledge, $A_1$’s choice of $p_{1,2}$ and $A_2$’s choice of $p_{2,1}$ are trivial, but with abilities being private knowledge, utility maximization by players $A_1$ and $A_2$ will ensure that information is processed exactly the same way as in a Bayesian equilibrium. Of course, there are other possible ways of how players could form their beliefs.

Note that in the model, we take a dual selfes perspective in which an individual is modeled as two separate players (e.g., worker 1 and player $A_1$), one for the individual’s rational decisions, the other one for the individual’s affective responses. Esteem is expressed by players $A_1$ and $A_2$ (i.e., $A_1$’s and $A_2$’s choices of $s_1$ and $s_2$), which imposes externalities on both workers’ utilities. Modeling each worker’s emotional system as an additional player who uses Bayesian updating allows us to apply existing solution concepts, like perfect Bayesian Nash equilibrium, to solve the

\[\text{footnoteindent}\text{footnotept}{footnotetext}{5With pure pride and $\delta_i = -s_i$, $\delta_j = -s_j$, the utility function (which is derived in the next section) can be rewritten as $u_i = e_1 + e_2 - \Phi(e_i/\theta_i) + f(\alpha \delta_i - \delta_j)$; revealing that $i$’s utility decreases in $j$’s pride and increases in her own pride. Hence, if $\alpha > 0$ the preference for pride is a status preference. A worker’s own contribution triggers pride, which is observed by others. A worker’s pride increases her own utility (similar to the “warm glow of giving” in Andreoni, 1989) but imposes a negative externality on the other worker’s utility.}\]
4.3 Workers Preferences for Esteem

To incorporate a demand for esteem into workers’ utility functions, we extend the utility functions introduced in section 3.1. Assume that workers’ preferences are given by

\[ u_i = \pi_i(e_1, e_2) - \Phi\left(\frac{e_i}{\theta_i}\right) + f(a_i) \]

where \( a_i = s_j - \alpha s_i \) gives the utility resulting from esteem. The parameter \( \alpha \in (0, 1) \) is the strength of the status preference (higher values imply stronger status effects). The function \( s_i(s_j) \) denotes esteem expressed by \( A_i(A_j) \), which depends on both workers’ relative efforts, as described in section 4.1. Substituting for \( a_i \), the utility function can be written as

\[ u_i(e_i, e_j, p_{1,2}, p_{2,1}) = e_i + e_j - \Phi\left(\frac{e_i}{\theta_i}\right) - \alpha s_i\left(\frac{e_i}{\theta_i}, \frac{e_j}{\theta_j}\right) \quad \text{for} \quad i, j \in \{1, 2\} \quad i \neq j. \]

The expression of esteem (described in section 4.1) is an evaluation of a worker’s choice of effort. We implicitly assume that esteem utility \( (a_i) \) is an anticipated emotion which is not experienced when choosing effort but it is expected to be experienced in at the end of stage three, after esteem is expressed. We assume that the expression of esteem takes place against the backdrop of specific organizational practices known to all workers, so that workers can anticipate esteem correctly. In other words, esteem is a component of the expected consequences of the choice of effort.

5 Abilities are Common Knowledge

Assume that, after abilities are drawn, all workers learn about abilities (cf. Figure 1). To solve the game, we find the subgame-perfect Nash equilibria of the extended game. With abilities being common knowledge, there are 4 plus 9 subgames. For each distribution of abilities there is a subgame. Denote these subgames by 1,2,3 and 4 (see Figure 1). In these four subgames abilities are given by \( (\theta^l, \theta^l) \), \( (\theta^l, \theta^h) \), \( (\theta^h, \theta^l) \), and \( (\theta^h, \theta^h) \). The additional 9 subgames are the subgames starting at the nodes \( E1 \) to \( E9 \), the subgames in which players \( A_1 \) and \( A_2 \) choose their affective responses. Using backward induction, we start by solving these subgames.

Assume that player \( A_1 \) (\( A_2 \)) has the same information as worker 1 (worker 2), which means that \( A_1 \) (\( A_2 \)) knows all workers’ abilities. Since \( A_1 \)’s and \( A_2 \)’s utilities only depend on their beliefs players \( A_1 \) and \( A_2 \) will choose the correct probabilities \( (p_{1,2} \text{ and } p_{2,1}) \) in each of the 9 subgames. The probabilities that maximize \( A_1 \)’s and \( A_2 \)’s utilities are
p_i = (0, 1, 1, 1, 1, 1, 0, 0)  p_2 = (0, 0, 0, 1, 1, 1, 1, 1)  (5)

where the k-th element of p_i corresponds to the utility-maximizing choice at node E_k. For subgames 1 to 4 the utility-maximizing beliefs are summarized in Table 2 where \( p^k_{i,j} \) denotes i’s expectation about j’s ability in subgame k.

For all subgames, the resulting in utilities are \( u_{A_1} = u_{A_2} = 0 \). By solving the subgames beginning at nodes E1 to E9 we determined the affective responses (which are \( A_1 \)’s and \( A_2 \)’s best-responses to all possible choices of effort). Equation 1 can be rewritten as

\[
s_i \left( \frac{e_i}{\theta_i}, \frac{e_j}{\theta_j} \right) = \beta_1 \frac{e_j}{\theta_j} - \beta_2 \frac{e_i}{\theta_i} \text{ for } i, j \in \{1, 2\}, \quad i \neq j.  \tag{6}
\]

Substituting \( s_i \) and \( s_j \) into the utility function, we can write a worker’s utility as a function of effort levels and abilities.

\[
u_i(e_i, e_j) = e_i + e_j - \Phi(e_i/\theta_i) + f \left[ \left( \frac{\beta_1 e_i}{\theta_i} - \beta_2 \frac{e_j}{\theta_j} \right) - \alpha \left( \beta_1 \frac{e_j}{\theta_j} - \beta_2 \frac{e_i}{\theta_i} \right) \right] \\
= e_i + e_j - \Phi(e_i/\theta_i) + f \left[ (\beta_1 + \alpha \beta_2) \frac{e_i}{\theta_i} - (\beta_2 + \alpha \beta_1) \frac{e_j}{\theta_j} \right] \\
= e_i + e_j - \Phi(e_i/\theta_i) + f \left[ \frac{\gamma_1 e_i}{\theta_i} - \gamma_2 \frac{e_j}{\theta_j} \right] \tag{7}
\]

where \( \gamma_1 = \beta_1 + \alpha \beta_2 \) and \( \gamma_2 = \beta_2 + \alpha \beta_1 \).

Next, we use \( A_1 \)’s and \( A_2 \)’s best-responses for each subgame (see Table 2) and proceed to workers’ choices of effort. We denote a worker’s strategy as \((X, Y)\), as before. The set \( S_i \) contains four pure strategies: \((e^l, e^l)\), \((e^h, e^l)\), \((e^l, e^h)\) and \((e^h, e^h)\). We are interested in pure strategy equilibria.

**Subgame 1:** Subgame is not a proper game. Both workers have low abilities, so by assumption \( e_i = e_j = \theta^l \) and \( \Phi(e_i/\theta_i) = c \) for \( i = 1, 2 \). Using \( u^k_i(e_i, e_j, p^k_{i,2}, p^k_{2,1}) \) to denote \( i \)'s utility in subgame
A worker's utility is given by

\[ u_i(e^l, e^l, p_{1,2}^{l,4}, p_{2,1}^{l,4}) = 2e^l - c + f(\gamma_1 - \gamma_2) \quad \text{for} \quad i = 1, 2. \]  

**Subgame 2:** In subgame 2, \( \theta_1 = \theta^l = e^l \) but \( \theta_2 = \theta^h \), so worker 2 can choose between high and low effort. Utilities are

\[ u_1^2(e^l, e^l, p_{1,2}^{l,2}, p_{2,1}^{l,2}) = 2e^l - c + f(\gamma_1 - \gamma_2) \frac{e^l}{\theta^h}, \]

\[ u_2^2(e^l, e^l, p_{1,2}^{l,2}, p_{2,1}^{l,2}) = 2e^l + f(\gamma_1 \frac{e^l}{\theta^h} - \gamma_2), \]

and

\[ u_1^2(e^h, e^h, p_{1,2}^{l,2}, p_{2,1}^{l,2}) = e^l + e^h - c + f(\gamma_1 - \gamma_2), \]

\[ u_2^2(e^h, e^h, p_{1,2}^{l,2}, p_{2,1}^{l,2}) = e^l + e^h - c + f(\gamma_1 - \gamma_2). \]

In subgame 2, worker 2 chooses high effort if \( u_2^2(e^l, e^h, p_{1,2}^{l,2}, p_{2,1}^{l,2}) \geq u_2^2(e^l, e^l, p_{1,2}^{l,2}, p_{2,1}^{l,2}), \) i.e., if

\[ c - e^h + e^l \leq f(\gamma_1 + \gamma_2) - f\left(\gamma_1 \frac{e^l}{\theta^h} - \gamma_2\right). \]  

**Subgame 3:** Subgame 3 is identical to subgame 2 except that the roles are reversed. Hence, worker 1 chooses high effort if equation 9 holds and the equilibrium strategy profile is given by \((e^h, e^l, p_{1,2}^{l,3}, p_{2,1}^{l,3})\).

**Subgame 4:** In subgame 4 utilities are

\[ u_1^4(e^l, e^l, p_{1,2}^{l,4}, p_{2,1}^{l,4}) = u_2^4(e^l, e^l, p_{1,2}^{l,4}, p_{2,1}^{l,4}) = 2e^l + f\left(\gamma_1 \frac{e^l}{\theta^h} - \gamma_2 \frac{e^l}{\theta^h}\right) \quad \text{for} \quad i = 1, 2 \]

if both workers choose low effort,

\[ u_1^4(e^l, e^h, p_{1,2}^{l,4}, p_{2,1}^{l,4}) = e^l + e^h + f\left(\gamma_1 \frac{e^l}{\theta^h} - \gamma_2 \right) \quad \text{and} \]

\[ u_2^4(e^l, e^h, p_{1,2}^{l,4}, p_{2,1}^{l,4}) = e^l + e^h - c + f\left(\gamma_1 - \gamma_2 \frac{e^l}{\theta^h}\right) \]

if 1 chooses low effort and 2 chooses high effort,

\[ u_1^4(e^h, e^l, p_{1,2}^{l,4}, p_{2,1}^{l,4}) = e^l + e^h - c + f\left(\gamma_1 - \gamma_2 \frac{e^l}{\theta^h}\right) \quad \text{and} \]

\[ u_2^4(e^h, e^l, p_{1,2}^{l,4}, p_{2,1}^{l,4}) = e^l + e^h + f\left(\gamma_1 \frac{e^l}{\theta^h} - \gamma_2\right). \]
if 2 chooses low effort and 1 chooses high effort, and
\[
u^i_1(e^h, e^h, p^1_{1,2}, p^2_{2,1}) = 2e^h - c + f(\gamma_1 - \gamma_2) \quad \text{for} \quad i = 1, 2
\]
if both workers choose high effort. Strategy profile \((e^l, e^l, p^1_{1,2}, p^2_{2,1})\) is an equilibrium in subgame 4 if \(u^i_1(e^l, e^l, p^1_{1,2}, p^2_{2,1}) \geq u^i_1(e^h, e^l, p^1_{1,2}, p^2_{2,1})\), i.e.,
\[
c - e^h + e^l \geq f \left( \gamma_1 - \gamma_2 \frac{e^l}{\bar{e}^h} \right) - f \left( \gamma_1 \frac{e^l}{\bar{e}^h} - \gamma_2 \frac{e^l}{\bar{e}^h} \right),
\]
and strategy profile \((e^h, e^h, p^1_{1,2}, p^2_{2,1})\) is an equilibrium if
\[
c - e^h + e^l \leq f(\gamma_1 - \gamma_2) - f \left( \gamma_1 \frac{e^l}{\bar{e}^h} - \gamma_2 \right).
\]
The asymmetric strategy profiles \((e^l, e^h, p^1_{1,2}, p^2_{2,1})\) and \((e^h, e^l, p^1_{1,2}, p^2_{2,1})\) are equilibria if \(u^i_1(e^l, e^h, p^1_{1,2}, p^2_{2,1}) \geq u^i_1(e^h, e^l, p^1_{1,2}, p^2_{2,1})\) and \(u^j_1(e^h, e^l, p^1_{1,2}, p^2_{2,1}) \geq u^j_1(e^l, e^h, p^1_{1,2}, p^2_{2,1})\) (for \(i, j \in \{1, 2\}, \ i \neq j\)), i.e.,
\[
f \left( \gamma_1 - \gamma_2 \frac{e^l}{\bar{e}^h} \right) - f \left( \gamma_1 \frac{e^l}{\bar{e}^h} - \gamma_2 \frac{e^l}{\bar{e}^h} \right) \geq c - e^h + e^l \geq f(\gamma_1 - \gamma_2) - f \left( \gamma_1 \frac{e^l}{\bar{e}^h} - \gamma_2 \right).
\]

5.1 Equilibria for the Proper Game

To simplify notation, we define
\[
\Delta f^i := f \left( \gamma_1 - \gamma_2 \frac{e^l}{\bar{e}^h} \right) - f \left( \gamma_1 \frac{e^l}{\bar{e}^h} - \gamma_2 \frac{e^l}{\bar{e}^h} \right) > 0,
\]
\[
\Delta f^h := f (\gamma_1 - \gamma_2) - f \left( \gamma_1 \frac{e^l}{\bar{e}^h} - \gamma_2 \right) > 0 \quad \text{and} \quad z := c - e^h + e^l > 0.
\]
\(\Delta f^i\) is the increase in utility from esteem if a worker increases her effort from \(e^l\) to \(e^h\) while the other worker exerts less than maximal effort. Similarly, \(\Delta f^h\) is the increase in utility from esteem if a worker increases her effort from \(e^l\) to \(e^h\) while the other worker exerts maximal effort, and \(z\) denotes the opportunity cost of increasing effort. Note that \(\Delta f^i \leq \Delta f^h\) is \(f\) is concave and \(\Delta f^i \geq \Delta f^h\) if \(f\) is convex.

In the game with common knowledge of abilities and for given abilities and costs of effort (i.e., for given parameters \(\theta^l, \theta^h, c\)) the set of equilibria depends on \(\gamma_1\) and the curvature of \(f\) and is completely characterized by equations 9 to 12. For concave and convex preferences over esteem, the equilibria are summarized in Table 3.

In principle, all three ways of expressing esteem (see section 4.1) can lead to workers choosing high effort. The upper part in Table 3 summarizes the equilibria if the preference for esteem is
Table 3: Equilibria in the game with esteem and complete information.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Equilibria</th>
</tr>
</thead>
<tbody>
<tr>
<td>$f'' &lt; 0$</td>
<td>$z \leq \Delta f^l \leq \Delta f^h$</td>
</tr>
<tr>
<td>(concave)</td>
<td>$\Delta f^l \leq z \leq \Delta f^h$</td>
</tr>
<tr>
<td></td>
<td>$\Delta f^l \leq \Delta f^h \leq z$</td>
</tr>
</tbody>
</table>

| $f'' > 0$ | $z \leq \Delta f^h \leq \Delta f^l$ | $\{((e^h, e^h), (e^h, e^h), p_1, p_2)\}$ |
| (convex)   | $\Delta f^h \leq z \leq \Delta f^l$ | $\{((e^l, e^l), (e^l, e^l), p_1, p_2), ((e^l, e^h), (e^l, e^l), p_1, p_2)\}$ |
|             | $\Delta f^h \leq \Delta f^l \leq z$ | $\{((e^l, e^l), (e^l, e^l), p_1, p_2)\}$ |

6 Abilities are Private Knowledge

Assume that after abilities are drawn, workers do not learn about others’ abilities. This implies that a worker’s type is private information. Due to the sequential structure of the game, esteem is expressed after both workers’ choices of effort are observed. A worker’s choice of effort can signal her ability. The expression of esteem by player $A_1$ depends on her conditional expectation about worker 2’s ability, $E_1[\theta_2|e_2]$, and is given by

$$s_1 \left( \frac{e_1}{\theta_1}, \frac{e_2}{E_1[\theta_2|e_2]} \right) = \beta_1 \frac{e_2}{E_1[\theta_2|e_2]} - \beta_2 \frac{e_1}{\theta_1}. $$
Similarly, we can write the expression of esteem by player $A_2$ as

$$s_2\left(\frac{e_2}{\theta_2}, \frac{e_1}{E_2[\theta_1|e_1]}\right) = \beta_1 \frac{e_1}{E_2[\theta_1|e_1]} - \beta_2 \frac{e_2}{\theta_2}.$$  

With incomplete information about abilities, worker 1 does not know $s_2$ because $A_1$ does not know worker 2’s ability ($\theta_2$) and $A_2$’s conditional expectation ($E_2[\theta_1|e_1]$). Taking expectations we write $A_1$’s expectation about esteem expressed by player $A_2$ as

$$s_2\left(\frac{e_2}{E_1[\theta_2]}, \frac{e_1}{E_1[\theta_2|e_1]}\right) = \beta_1 \frac{e_1}{E_1[\theta_2|e_1]} - \beta_2 \frac{e_2}{E_1[\theta_2|e_2]}$$  

where $E_1[\theta_1|e_1]$ is $A_1$’s second-order belief about worker 1’s ability. We assume that $E_1[\theta_2|e_1] = E_2[\theta_1|e_1]$.

As above, we assume that player $A_1$ ($A_2$) has the same information as worker 1 (worker 2). Players $A_1$ and $A_2$ cannot distinguish between all nodes $E_1 - E_9$ because some of these nodes belong to the same information set. The information sets that players $A_1$ and $A_2$ can distinguish are given by

\[
\{(E1, E2), (E3), (E4), (E5, E8), (E6), (E7, E9)\} \quad \text{for player } A_1 \quad \text{and} \\
\{(E1, E9), (E8), (E4), (E3, E6), (E5), (E2, E7)\} \quad \text{for player } A_2.
\]

$A_1$ knows worker 1’s ability (by assumption) and knows worker 2’s ability only if worker 2 chooses high effort. For information sets that are singletons, a player knows both workers’ types for sure. At node $E_3$, for example, player $A_1$ knows worker 1’s type by assumption, and worker 2’s type because by choosing high effort, worker 2 reveals her true type. This implies that players $A_1$ and $A_2$ maximize their utilities by choosing

$$p_{1,2}(E3) = p_{1,2}(E4) = p_{1,2}(E6) = 1$$  

and

$$p_{2,1}(E4) = p_{2,1}(E5) = p_{2,1}(E8) = 1.$$  

These probabilities determine $A_1$’s ($A_2$’s) expectation about worker 2’s (worker 1’s) type and determine the expression of esteem (according to equation 13) at the information sets that are singletons. The utility-maximizing choices of players $A_1$ and $A_2$, conditional expectations and the expression of esteem at information sets that are not singletons depend on the equilibrium strategy profile. In the following section, we look at the decision of high ability workers who can choose high or low effort, i.e., the set $S_2$ contains two pure strategies, $S_2 = \{e^l, e^h\}$. Since conditional expectations depend on the equilibrium strategy profile, we consider the different perfect Bayesian equilibria in turn.
6.1 Pooling Equilibrium

Assume that in equilibrium both workers choose low efforts, regardless of their abilities. This is a pooling equilibrium because effort choices reveal no information about workers’ abilities. At all \(E\)-nodes that are reached in equilibrium (\(E1, E2, E7, E9\)), players \(A1\) (\(A2\)) only knows worker 1’s (worker 2’s) type. Maximizing their utilities (equations 2 and 3), \(A1\) and \(A2\) will choose

\[
p_{1,2}(E1) = p_{1,2}(E2) = p_{1,2}(E7) = p_{1,2}(E9) = p
\]

and

\[
p_{2,1}(E1) = p_{2,1}(E2) = p_{2,1}(E7) = p_{2,1}(E9) = p.
\]

Also, if worker 1 unilaterally deviates by choosing high effort, \(A1\)’s belief about worker 2’s type will be given by \(p_{1,2,1}(E1) = p\) but \(A2\)’s belief will be \(p_{2,1,2}(E1) = 1\). Similarly for deviations for worker 2, hence

\[
p_{1,2}(E5) = p_{1,2}(E8) = p \quad p_{2,1}(E5) = p_{2,1}(E8) = 1 \quad \text{if only worker 1 deviates, and}
\]

\[
p_{2,1}(E3) = p_{2,1}(E6) = p \quad p_{1,2}(E3) = p_{1,2}(E6) = 1 \quad \text{if only worker 2 deviates, and}
\]

\[
p_{2,1}(E4) = p_{1,2}(E4) = 1 \quad \text{if both workers deviate.}
\]

Using vector notation, we summarize these beliefs as

\[
\vec{p}_{1,2} = (p, p, 1, 1, p, 1, p, p) \quad \text{and} \quad \vec{p}_{2,1} = (p, p, p, 1, 1, p, 1, p). \tag{16}
\]

Given \(A1\)’s and \(A2\)’s beliefs, conditional expectations are given by \(E_i[\theta_j|e_l] = p_{i,j} \theta^h + (1-p_{i,j}) \theta^l\) for \(i, j \in \{1, 2\}, i \neq j\). For given conditional expectations, worker 1’s esteem utility is given by

\[
a_1(e_1, e_2) = \left( \beta_1 e_1 - \frac{\beta_2 e_2}{E_1[\theta_2|e_1]} - \alpha \beta_1 \frac{e_2}{E_1[\theta_2|e_1]} - \beta_2 \frac{e_1}{\theta_1} \right)
\]

\[
= \beta_1 e_1 - \gamma_2 \frac{e_2}{E_1[\theta_2|e_1]} + \alpha \beta_2 \frac{e_1}{\theta_1}. \tag{18}
\]

Note that the last term in equation 18 reveals that expected esteem utility depends on a worker’s own ability. In equilibrium, equation 18 can be rewritten as

\[
a_1(e^l, e^l) = \beta_1 e^l - \gamma_2 e^l + \alpha \beta_2 \frac{e^l}{\theta_1}. \tag{19}
\]

Assuming that worker 1 has high ability, her expected utility is given by
\[ u_1(e^l, e^l) = 2e^l + f \left( (\beta_1 - \gamma_2) \frac{e^l}{\theta} + \alpha \beta_2 \frac{e^l}{\theta} \right). \]

She can deviate by increasing her effort to \( e^h \). Because of the deviation, a different node is reached. If worker 2 has low ability, node \( E_8 \) is reached instead of node \( E_9 \). If worker 2 has high ability, node \( E_5 \) is reached instead of node \( E_7 \). At both nodes, worker 2 knows for sure that worker 1 has high ability. Hence, worker 2’s conditional expectation about worker 1’s ability is equal to her true ability, \( E_2[\theta_1 | e_1 = e^h] = \theta^h \). This changes esteem utility to

\[ a_1(e^h, e^l) = \beta_1 + \alpha \beta_2 - \gamma_2 \frac{e^l}{\theta} \]
and yields utility

\[ u_1(e^h, e^l) = e^l + e^h - c + f \left( \beta_1 + \alpha \beta_2 - \gamma_2 \frac{e^l}{\theta} \right). \]

Hence, \((e^l, e^l, \tilde{p}_{1,2}, \tilde{p}_{2,1})\) is a perfect Bayesian equilibrium if \( u_1(e^l, e^l) \geq u_1(e^h, e^l) \), i.e.,

\[ c - e^h + e^l \geq f \left( \beta_1 + \alpha \beta_2 - \gamma_2 \frac{e^l}{\theta} \right) - f \left( (\beta_1 - \gamma_2) \frac{e^l}{\theta} + \alpha \beta_2 \frac{e^l}{\theta} \right). \]  \hspace{1cm} (20)

### 6.2 Separating Equilibrium

Assume that \((e^h, e^h)\) is the equilibrium strategy profile. We call this the separating equilibrium because in equilibrium, low ability workers choose low efforts and high ability workers choose high efforts. In this separating equilibrium \( A_1 \)’s and \( A_2 \)’s beliefs are given by

\[ p_{1,2}(E1) = p_{1,2}(E2) = p_{1,2}(E5) = p_{1,2}(E7) = p_{1,2}(E8) = p_{1,2}(E9) = 0 \quad \text{and} \]

\[ p_{2,1}(E1) = p_{2,1}(E2) = p_{2,1}(E3) = p_{2,1}(E6) = p_{2,1}(E7) = p_{2,1}(E9) = 0 \]

and by equations 14 and 15. These probabilities maximize \( A_1 \)’s and \( A_2 \)’s utilities and can be summarized as

\[ \tilde{p}_{1,2} = (0, 0, 1, 1, 0, 1, 0, 0, 0) \quad \text{and} \]

\[ \tilde{p}_{2,1} = (0, 0, 0, 1, 1, 0, 0, 1, 0). \] \hspace{1cm} (21)

It follows that one worker’s conditional expectation about the other worker’s ability is given by the other worker’s effort, i.e.,
\[ E_i[\theta_j | e_j = \theta^i] = \theta^i \quad \text{and} \quad E_i[\theta_j | e_j = \theta^h] = \theta^h \quad \text{for} \quad i, j \in \{1, 2\} \quad i \neq j. \]  

This implies that equation 18 simplifies to

\[ a_1(e_1, e_2) = \beta_1 - \gamma_2 + \alpha \beta_2 \quad \text{for} \quad e_1, e_2 \in \{e^l, e^h\}. \]  

Assuming that worker 1 has high ability, her expected utility is given by

\[ u_1(e^h, e^h) = e^h - c + p e^h + (1 - p) e^l + f(\beta_1 - \gamma_2 + \alpha \beta_2). \]

Here, \( e^h - c \) is the net utility from worker 1’s effort and \( p e^h + (1 - p) e^l \) is the expected utility from the other worker’s effort. Worker 1 can deviate by decreasing her effort to \( e^l \), which yields utility

\[ u_1(e^l, e^h) = e^l + p e^h + (1 - p) e^l + f(\beta_1 - \gamma_2 + \alpha \beta_2\frac{e^l}{\hat{p}^n}). \]

Note that a high ability worker achieves a higher monetary payoff (\( 0 < c - e^h + e^l \)) but lower esteem utility as long as \( \beta_2 \) is positive. This is the guilt-effect: Although the other worker’s conditional expectation about 1’s ability is independent from 1’s actual effort (observing \( c_1 \) is not informative), esteem utility is lower because the strength of 1’s own affective response is inversely related to her own ability, which she knows (see the last term in eq. 18). Put bluntly, a deviation decreases 1’s esteem utility because she knows that she could have chosen higher effort.\(^6\)

The strategy profile \((e^h, e^h, \hat{p}_{1,2}, \hat{p}_{2,1})\) is a perfect Bayesian equilibrium if

\[ f(\beta_1 - \gamma_2 + \alpha \beta_2) - f\left(\beta_1 - \gamma_2 + \alpha \beta_2\frac{e^l}{\hat{p}^n}\right) \geq c - e^h + e^l. \]  

### 6.3 Characterization of Equilibria with Private Knowledge

In the game with incomplete information of abilities and for given abilities and costs of effort (i.e., given parameters \( \theta^l, \theta^h, c \)) the set of equilibria is completely characterized by equations 20, 25 and the corresponding beliefs (equations 16 and 17, and equations 21 and 22). Note that the beliefs represent the equilibrium strategy profiles for players \( A_1 \) and \( A_2 \), and, at the same time, they are the posterior beliefs which are computed according to Bayes rule. The equilibria are summarized in Table 4.

---

\(^6\)Or more precisely: \( s_2 \) remains constant but \( s_1 \) increases because 1’s affective response is stronger. In other words, if 1 deviates, 2 receives more esteem. This is because 1 knows that her own relative effort is smaller than 1 and takes this into account when expressing esteem. Compared to her own choice of relative effort, 2 has chosen higher relative effort.
Table 4: Equilibria in the game with esteem and incomplete information.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Equilibria</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Delta \tilde{f}^h \leq \Delta \tilde{f}^l$</td>
<td>$(\tilde{e}, \tilde{e}, \tilde{p}<em>{1,2}, \tilde{p}</em>{2,1})$ (pooling)</td>
</tr>
<tr>
<td>$\Delta \tilde{f}^h \leq \Delta \tilde{f}^l \leq z$</td>
<td>$(\tilde{e}, \tilde{e}, \tilde{p}<em>{1,2}, \tilde{p}</em>{2,1})$ (pooling)</td>
</tr>
<tr>
<td>(Case 1)</td>
<td>no equil. in pure strategies</td>
</tr>
<tr>
<td>$z \leq \Delta \tilde{f}^h \leq \Delta \tilde{f}^l$</td>
<td>$(\tilde{e}, \tilde{e}, \tilde{p}<em>{1,2}, \tilde{p}</em>{2,1})$ (pooling)</td>
</tr>
<tr>
<td>$\Delta \tilde{f}^h \geq \Delta \tilde{f}^l$</td>
<td>$(\tilde{e}, \tilde{e}, \tilde{p}<em>{1,2}, \tilde{p}</em>{2,1})$ (pooling)</td>
</tr>
<tr>
<td>(Case 2)</td>
<td>$(\tilde{e}, \tilde{e}, \tilde{p}<em>{1,2}, \tilde{p}</em>{2,1})$ (pooling and sep.)</td>
</tr>
<tr>
<td>$\Delta \tilde{f}^l \leq z \leq \Delta \tilde{f}^h$</td>
<td>$(\tilde{e}, \tilde{e}, \tilde{p}<em>{1,2}, \tilde{p}</em>{2,1})$ (pooling and sep.)</td>
</tr>
<tr>
<td>$z \leq \Delta \tilde{f}^l \leq \Delta \tilde{f}^h$</td>
<td>$(\tilde{e}, \tilde{e}, \tilde{p}<em>{1,2}, \tilde{p}</em>{2,1})$ (pooling and sep.)</td>
</tr>
</tbody>
</table>

With incomplete information about abilities we have to distinguish whether $\Delta \tilde{f}^h \leq \Delta \tilde{f}^l$ or $\Delta \tilde{f}^h \geq \Delta \tilde{f}^l$. The upper part in Table 4 summarizes the equilibria if $\Delta \tilde{f}^h \leq \Delta \tilde{f}^l$ (Case 1), the lower part summarizes the equilibria if $\Delta \tilde{f}^h \geq \Delta \tilde{f}^l$ (Case 2).

Case 1: If $\Delta \tilde{f}^h \leq \Delta \tilde{f}^l$ there is a unique equilibrium in pure strategies for high and for low values of $z$. For intermediate values of $z$ there exists no equilibrium in pure strategies. Note that $\Delta \tilde{f}^h = 0$ but $z > 0$ and $\Delta \tilde{f}^l > 0$ if $\alpha = 0$ or if $\beta_2 = 0$, which means that “pure respect” (see Section 4.1) cannot solve the problem of team production.

Case 2: If $\Delta \tilde{f}^h \geq \Delta \tilde{f}^l$ there is a unique equilibrium in pure strategies for high and for low values of $z$. However, for intermediate values of $z$ there are two equilibria in pure strategies.

In both cases, an equilibrium in high contributions exists only if the social comparison condition is fulfilled ($\beta_2 = 0$) so that social comparison matters for the expression of esteem, and if the status condition is fulfilled ($\alpha = 0$) so that esteem utility increases with the difference of esteem one receives and esteem received by others.

7 Results and Discussion

Within this paper, we make two contributions: On a theoretical level, we show how to incorporate affective responses into game theoretical models, and on a more applied level, we present a model illustrating how one particular affective response, the expression of esteem, can solve the problem of team production.

When workers’ efforts are not enforceable, esteem can lead to higher effort choices. The interaction of workers generates information about effort decisions on the basis of which workers evaluate their coworkers’ effort decisions (e.g., expressing esteem through peer feedback or performance appraisals). These evaluations can in turn influence effort decisions.
In teams whose composition is constant, it is likely that over time, workers learn about each others’ abilities, so that information about abilities is common knowledge. In teams whose members know each others abilities, esteem incentives, if strong enough, lead to optimal effort choices, even if social comparison does not matter for expression of esteem and even if the preference for esteem is not a status preference. To achieve optimal effort levels, organizations should create opportunities that ensure the expression of esteem, for example, by creating a work environment in which there is enough room for the exchange of esteem services, like paying attention, expressing your opinion, or giving credit (Brennan and Pettit, 2000, 89-90).

If workers are regrouped frequently into new teams, it is plausible that workers do not know each others’ abilities. Then, esteem incentives can lead to optimal effort choices only if the social comparison condition and the status condition are fulfilled. When both conditions are fulfilled, esteem is expressed for a worker’s effort relative to the effort chosen by a reference group and the preference for esteem is a status preference. To achieve optimal effort levels, organizations should not only encourage the expression of esteem. In addition, organizations should ensure that the expression of esteem is transparent, in the sense that workers are not only informed about their own evaluation but also about the evaluations of other team members, for example, if awards are given publicly (Neckermann and Frey, 2013).

On a theoretical level, our model illustrates how the role of emotions can be analyzed in game-theoretic models. By modeling each player’s emotional system as an additional player we are able to investigate affective responses while using existing game-theoretic analysis. Thus, we relax the unrealistic assumption about the irrelevance of emotions while maintaining formal rigor of rational choice modeling. When models based on simple motivational assumptions (e.g., payoff maximization) fail, an extended rational choice model as the one presented in this paper, provides an alternative to describe human behavior.
References


Neckermann, Susanne and Bruno S Frey, “And the winner is...? The motivating power of employee awards,” Journal of Socio-Economics, October 2013, 46, 66–77.


