University of Bath

PHD

Relational Team Contract and Inequity Aversion

Li, Yihang

Award date:
2018

Awarding institution:
University of Bath

Link to publication

Alternative formats
If you require this document in an alternative format, please contact:
openaccess@bath.ac.uk

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Download date: 27. Sep. 2021

Relational team contract and inequity aversion

Yihang Li

A report submitted in fulfilment of the requirements for the degree of Doctor of Philosophy

Univeristy of Bath

Department of economics

January 2018

COPYRIGHT

Attention is drawn to the fact that copyright of this thesis/portfolio rests with the author and copyright of any previously published materials included may rest with third parties. A copy of this thesis/portfolio has been supplied on condition that anyone who consults it understands that they must not copy it or use material from it except as permitted by law or with the consent of the author or other copyright owners, as applicable.

Abstract

The Thesis studies moral hazard problem in teams. We consider team production models where incentive can be provided through explicit sharing and relational contract. Incentive provision is discussed under various information structures. Under deterministic production, the output itself act as a strong signal of deterring shirking, thus noisy soft signal would not help to improve efficiency. While in cases where soft information is sufficient to infer agents' effort, we show that the optimal sharing would concentrate negative sharing on one agent who will be fully motivated by relational incentive.

We further studied a model with sub-teams where sub-team performances are deterministic signals to sub-team members' effort. The value of the certainty in the soft information crucially depends on the structure of the sub-teams. Once we can ensure some heterogeneity between the organizational structure within the sub-teams, strongest relational incentive can be provided to cope with unilateral deviation. A necessary and sufficient condition for implementing a target effort level is then provided under general sharing. However, once we restrict the sharing rule to be linear, utilities among agents are no longer transferable. We show that linear sharing can implement the efficient effort, but with more restrictions on surplus distributions among agents. In general, linear sharing can be applied without loss of generality only if the surplus distribution is relatively balanced.

Finally we had some preliminary discussion of non-monetary incentive provision based on inequality aversion model on linear public good games. Applying Fehr and Schmidt [1999]'s model with Bayesian game technique, we explore a boarder range of equilibriums with positive contributions. However, equilibrium behavior relies on how inequality is defined among players, future theoretical and experimental work needs to be done to enable inequality aversion as a tool of incentive provision.

Acknowledgements

I would like to express my special appreciation and thanks to my superviosr Prof. Shasi Nandeibam, you have been a tremendous mentor for me. I would like to thank you for encouraging my research and all the effort you spent on guiding me through the four-year research period. I would also like to thank all the other staffs in economic department who provided great support to my research.

A special thanks to my family. Words cannot express how grateful I am to my parents, grandparents and for all of the sacrifices that you?ve made on my behalf. I wouldn't be able to get that far without you.

I would also like to thank all of my friends who supported me in writing, and incented me to strive towards my goal.

Contents

Chapter 1

Introduction

Numerous firms constitute a corporate structure as such that is engineered to ensure the primary interests of shareholders. Over the past few decades, the conventional notion that a firm's ultimate goal is to maximize share value for its shareholders has dominated economic thought and processes. However, the detrimental consequences instigated by the 2008 financial crisis, prompted a serious reconsideration of these ideals, instead proposing an alternative economic philosophy. Other issues were brought into the spotlight, for instance the maximization of shareholder value may incentivize shareholders to adopt a greater risk decision making process, which would consequently further intensify the insecurities of debtors, taxpayers and the wider economy as a whole. As a result, in response to the model of shareholder primacy, this dissertation attempts to seek an alternative approach in the form of "team production" which many scholars consider of fundamental importance in understanding how corporate self-interested objectives can generate social and economic spillovers.

Team incentive problem has long been studied, but lack progress compared to other areas in incentive theory. In the early literature, team production is defined as " production in which (1)several resources are used; (2) the product is not a sum of separable outputs of each cooperating resource; (3) not all resources used in team production belong to one person." The above characteristics form the basic benchmark for all later studies of team production, although sometimes people relax the second condition and assume linear production technology for simplicity. Unlike the principal-agent problem which strongly emphasis the share value (the principal's benefit), the team production problem deals with a more symmetric relationship. Each team member is expected to exert effort while the residual claimant is not pre-attributed to anyone. Thus, the principal-agent problem can be treated as a special case of the team production problem where (in most cases) the principal does not need to exert effort and claims all the residuals, in the mean time the agent is hired by the principal who gets a conditional wage based on the produced output.

Moral hazard arises when effort cannot be directly contracted upon. When team members' payoff can only contingent on the final output, i.e the contracted sharing rule, there will be incentive for some team members to free ride on other's effort(Holmstrom [1982]). Holmstrom [1982] also shows that if the output will always be fully distributed among team members, efficient outcome can never be achieved. Intuitively, each team member only receive a fraction of the total output, any team member will not be fully benefitted from an increase of effort from himself nor will he be punished fully from an decrease of effort from himself only.

A solution provided by Holmstrom introduces the idea of budget breaking. In the solution, efficient outcome is achieved through a scheme such that the whole team will get punished if the output is under certain threshold. This requires the whole team giving away some surplus though some mechanism. To overcome the enforcement problem of such contract, Holmstrom [1982] introduces a special role into the team setting - the "principal" or budget breaker whose job is to break the budget and punishing the whole team when the output is low.

Another idea investigated by McAfee and McMillan [1991] looked the case the other way round: Instead of breaking budget through punishing the whole group ex-post, each agent would be receiving 100% share of the production and the principal balances the budget through taking fixed payments from agents ex-ante. According to McAfee and McMillan [1991]'s solution, the principal can merely act as a third party money keeper and should not be able to influence the production by any means.

Both of the above contracts work based on the fact that they provide strongest incentive possible to agents, on the other hand budget balancing shift all the negative incentives to the principal who in theory don't exert effort and thus have no incentive problem. These two models indeed provided a neat solution towards moral hazard problem in teams, but the role of the principal has suffered lots of criticisms. As the principal does not involve in the production himself and can only get positive payoff when the output is low. Eswaran and Kotwal [1984] shows it may give the principal incentive to sign some side-contracts with one or more team members making them benefit from producing a low output.

When the special principal is removed from the team, most of the nice properties of the optimal contract would collapse. To improve efficiency, the contract has to utilize more information other than the final output level. Without doubt, when a group of people working together, they should have a general idea on how hard working each others are. These kind of information, although might contain noise, can be used in contracting as long as they bring in useful information on agents' effort. However, such information, if written in contracts, will face a crucial obstacle in the implementation, that is the verifiability of such information.

In most cases, the additional information can be observed by team members only. Thus writing a contract with unverifiable information would signify that no third party will be able to enforce the contractual parties to follow what the contract specifies. Just as how meaningless it is to condition the contract on a variable whose value cannot be objectively evaluated. However, this problem has recently been tackled by the so called relational contract, which tries to take advantage of soft information that cannot be verified and eliminate the necessity of the third party enforcement.

In reality, most contractual relationships are long lasting, making the incentive provision more complicated than in the static setting. The well known folk theorem states that if a game is played infinitely repeatedly, for large enough discounting factors, any outcome that can gives players' payoffs larger then their minmax payoffs can be sustained as a Nash equilibrium. In a way we know what can be achieved but folk theorem does not answer how can it be done. In infinitely repeated games, the threat of falling back to a "bad" equilibrium will give motivations to players to keep on the "good" equilibrium path. When applied to contract theory, the above idea forms the bases of implicit contracting which is also called relational contract (Levin [2003]).

Although contracts only specify the rights and obligations of the contractual parties, the effectiveness of the contracts must rely on the ability of a third party, often the court, to enforce them. Relational contract studies the enforcement problem in incentive theory, pointing out that the court might not be able to enforce every agreement in a contract. For example in a team setting, each team member may have a better idea of everyone's contribution than the court does, writing explicit contracts with respect to the contribution of each individual may lead to distortion in incentive provision. It is often observed that, the payments in team contracts are only conditioned on the final output, while any bonus or punishment is done through the inner regulation system in which the third party is not involved. The infinite repeated game framework enables firms to condition payments on soft information that cannot be verified by the court. However certain conditions must be met to make sure the implicit understanding between contractual parties will indeed be executed.

The growing focuses on relational contracts shed light on team moral hazard problem. The enforcement problem in the static game can be mitigated by relational contract while the principal is not a must to break the budget. Moreover, more complicated contracts can be utilized providing more possibilities in getting closer to the first best outcome, without need to consider the enforcement ability of the court in reality.

While relational contract provides another direction in studying traditional team moral hazard problem, some recent empirical and experimental findings should not be neglected. The standard economic theory was developed under an implicit assumption that individuals only care about their own interest, on the other hand Fehr and Schmidt [1999] shows it's not always the case. Human beings are not machines that work for a purely pre-determined purpose, sometimes individuals' decision making involves in emotional factors that are not counted into classical economic theory.

It has been consistently observed in different experiment that people dislike inequality. Such preference, named inequity aversion, means that people will feel envy when getting less than their companions while feel guilty when earning more than their colleagues. Fehr and Schmidt [1999] models inequity aversion and builds into the principal-agent model. However, we may doubt that in an asymmetric relationship between principal and agent, would the principal's payoff be observable to the agent? Particularly if the incentive provision under inequity aversion would be costly, the principal may mask his payoff and leave it unobservable to the agent.

We find it may be more appropriate to study inequity aversion under symmetric relations. Moreover, surveys of employee opinions within firms have shown that employees compare not only relative salaries, but also relative performance against that of co-workers. To account for inequity aversion, there is need to build not only relative payoffs but also relative performance in individuals' utility functions.

The three topics we described above represents the main themes we are going to analyze in this dissertation. However, the theory of inequality aversion are developed based on basic experiments i.e. dictator game, public good games. Whether the existing theory can be applied to more complicated circumstances is unknown. Before we develop any team-relational contract theory with inequality aversion, more theoretical and experimental work needs to be done. Thus in this dissertation we would use two chapters to analyze relational team contract and a separate chapter as a preliminary discussion on inequality aversion theory and its implication for future study.

Relational contract in teams

Rayo [2007]'s model provides us a benchmark to analyze relational team contract. We would start with similar settings and investigate some interesting areas that remain to be explored. We would first follow Rayo [2007]'s path supposing the team maximize its total surplus, while adding two different variations in the assumptions that we are interested in. We first consider the case with deterministic production with stochastic individual performance signals, to see if the additional signals could add anything to the model. Typically we will compare trigger strategy that depends only on the deterministic output with relational contract that can depend on both hard and soft information. We find that aggregate hard information would be sufficient to obtain the first best outcome, while noisy soft information would not add anything more. The second aspect we want to explore is what the linear assumption on the sharing rule plays in the model. Due to its simplicity, linear sharing is widely used in partnership and team contracts. But can such simple sharing be optimal? Is the practical advantage of linear sharing is at the cost of efficiency? Such questions need to be looked into more carefully. We would follow Rayo [2007]'s discussion under non-deterministic production, while we no longer put any restriction to the sharing rule. The results show that using linear sharing rule is without loss of generality when first-order approach is valid, however the optimal sharing structure might vary if we allow negative sharing. All these discussions will be presented in chapter 3.

Inspired by the value of certainty we have found in chapter 3, we asked ourselves another question: is there a way we can get other soft information without noise? Apparently it's very unrealistic to assume we can get perfect effort signal under individual level. But we will be able to construct sub-teams among the agents and get deterministic sub-team performances Nandeibam [2002]. In chapter 4 we will study a new relational team contract model with sub-team structure. We are typically interested in the following issues:

- How can we design the organizational structure of the team such that the sub-team performances can be utilized to the maximum in relational contracting.
- The optimal relational incentive under the previous sub-team structure.
- The efficiency that can be achieved under the combination of explicit sharing and relational incentive.

- The optimal linear sharing rule.
- The comparison of linear sharing with general sharing and the limitations of linear sharing

The most striking finding is that, once we restrict the sharing rule to be linear, the utilities among agents will no longer be transferable. Namely we cannot freely distribute team's surplus with linear sharing. Thus, what Rayo [2007] has been solved under linear sharing rule is only a special case with utilitarian social welfare function. It would be very interesting to see how general sharing is different from linear sharing and to what extend. We will spend some paragraphs together with a numerical example to have a detailed discussion over this issue in chapter 4.

Inequality aversion

Among all the experiments over other regarding preferences, linear public good is the one that is most similar to team production problem. Our theory model would be based on Fehr and Schmidt [1999] and Fehr et al. [1997]'s model of linear public good experiments. However, we point out two assumptions that are not very realistic under their experimental framework and try to develop new theoretical insights over the model. First assumption Fehr and Schmidt [1999] applies is that players' preferences over inequality are publicly known to everyone. This is a very strong assumption especially under experimental circumstances where subjects don't know who they are cooperating with. In chapter 5 we would assume that the parameters representing agents' preferences towards inequality are private information. But the distribution of each parameter are independent and identically distributed across all players and publicly observable. Another assumption or approach used by Fehr and Schmidt [1999] is that inequality aversion is calculated based on individual contribution levels even when the players' contribution vector will not be revealed. Thus we tried to redefine the inequality aversion based on the public information and see how the equilibrium behavior would differ. As a preliminary work, we will also point out the direction for future studies, especially the experimental variations we are interested in.

As discussed above, team moral hazard problem, relational contract and inequity aversion form the basic elements in this dissertation. Although the theory has been developed for some time, there are still wide ranges of issues remain uninvestigated. A study on recent literature also indicates a need to combine all the three elements together to conduct a research in team incentive problem, which is the main objective of this dissertation. The content of thesis is arranged with the following order: the second chapter will review all the relevant literature; the third chapter which act as a foundational chapter will formally introduce the idea of team relational contract based on Rayo [2007]'s model with several extensions; The fourth chapter where our main results lie, will have a detailed discussion over the team organization structure, efficiency and sharing rules with relational contract; the fifth chapter is a preliminary theoretical chapter for future experimental studies, we will analyze Fehr and Schmidt [1999]'s inequality aversion model with Bayesian game approach; the last chapter summarizes our findings.

Chapter 2

Literature Review

This chapter summarizes the literature[1], both classical and recent ones, on moral hazard in teams, relational contract and inequity aversion.

In a simple model with one principal and one agent, we know that moral hazard problem is solved by conditioning the agent's wage on the final output. When the principal does not exert effort, the efficient outcome for the economy as a whole can be maintained at a cost from the principal side. However, in team production where every team members are supposed to contribute towards the final output, incentive provision may be very hard and tedious, especially when the total output generated by the team is the only information that can be measured and contracted upon. In this case, free rider problem arises when the share of output is the only source of incentive (Holmstrom [1982]).

[1] Here we only include general review of the relative literatures, for more discussion on detailed topics please refer to each chapters respectively.

When team members' payoffs are expressed as functions of the output, we describe the set of functions across all members as the sharing rule within the team. A sharing rule is balanced if the share received by each agent can be summed up to the final output, namely the output will be fully distributed among the team members under balanced sharing. Holmstrom [1982] shows that under balanced sharing rules, each agent will be willing to exert first-best effort if and only if the agent received a full share of the output. Otherwise agents' would always have incentive to shirk. As the efficiency loss by shirking will be shared among all team members while the agent who shirks saves the full effort cost.

Therefore Holmstrom [1982] suggested a punishing scheme which solves the incentive problem by introducing a principal who doesn't exert effort into the team. No matter who shirks, the principal will be able to punish the entire group based on the output level. However, this heavy punishment makes the principal benefit from project failure which makes him prefer the agents to shirk. Thus the enforcement of such budget breaking scheme has been doubted. The moral hazard problem arises on the principal's side that he can collude with one agent making him shirk(Eswaran and Kotwal [1984]).

Apart from the budget-breaking scheme, relative performance measure or tournament is widely used. In firms that use internal promotion as incentives, the principal could create a pool of rewards, and each agent's reward would be purely based on his performance ranking. Since the total payment from the principal is fixed, the principal can no longer take advantage of information asymmetry to make himself better off in this case. Moreover, Green and Stockey [1983] shows that this mechanism can be very close to optimal when the number of workers doing the same job becomes sufficiently large, the ranks of the workers become an effective and accurate evaluation of their effort.

An obvious shortcoming of relative performance evaluation is that it introduces competition among agents and discourages cooperation (Mookherjee [1984]). Itoh [1991] studied a slightly different model where the agents' tasks are affected by independent and privately observable noises but they can help other workers accomplishing their tasks. If each agent's wage depends only on how well he can accomplish his own task, he will not spend effort to help others since it's costly to him while bringing in no benefit. But the principal can use joint performance evaluation and create conditions to get the agents help each other.

The above results are established in a static framework such that no future interaction between the contractual parties are considered. However, in reality, employment relationship and partnerships are often long lasting which makes relational incentives a possibility in incentive provision. In other words, people's decisions are not only affected by current payoffs but also by the future interest, i.e. reputation. The long-lasting nature of contract and infinite-repeated game framework serves as the benchmark of self-enforcing contracts or relational contract, where incentives are provided through the future gains from continuing the relationship.

The early literature on incentive theory focuses on asymmetric information where principal is not able to observe certain characteristic of the agent. Through designing a contract contingent only on the observable variables, the principal can give the agent some incentive to do what the principal desires at a cost which must be taken into account. There is an important implicit assumption that the contract once signed, there will be a third party to enforce both contractual parties to practice their commitment. However, in the real world the enforcement relies heavily on the verification ability of the court. In other words, although something can be observed by the contracting parties, it may not be possible for a third party to verify it, or the verification is costly. In this case, the contract has to involve some implicit understanding that must be self-enforcing.

In the principal-agent relational contract model, a principal employs an agent who chooses a costly effort level (action) which is positively correlated with the expected output. Neither output nor effort the agent spent can be verified by a third party. Therefore, an explicit contract cannot be signed since the court is unable to verify it. In such case, can any bonus or punishment be exercised in order to give incentive to the agent? The answer is yes. The relational contract specifies a discretionary incentive payment either from the principal or from the agent, together with the outcome if anyone reneges on the discretionary payment. The threat of getting into a bad outcome onwards helps to enforce the execution of such payment. Levin [2003] introduces a self-enforcing constraint under which incentive can be implemented. The so called self-enforcing constraint says that if the agent exert the target effort level, the principal must be willing to pay any bonus to the agent and in the meantime the agent must be willing to pay any penalty to the principal. This constraint also gives the limitation on the incentive can be provided.

Assuming risk neutrality, optimal relational contract has the nice property of being stationary (Levin [2003]). In the moral hazard model in Levin [2003], where the agent's performance is not perfectly objectively measurable, the form of optimal contract is rather monotonous: the principal sets a fixed payment and adjusts this base payment according to the output. If output exceeds some level, he pays more to the agent and if output is below this level, he pays less. Risk neutrality allows the principal to set the strongest incentive. However, self-enforcement itself sets the boundary of punishment and reward. The reason is that, the value generated from continuation of the relationship is limited, thus the reward and punishment to the agent is limited.

Parties cannot contract on the imperfect objective performance measures, even though in a relational contract they can rely on self-enforcement to agree with them. In practice, subjective measurement is widely used to alleviate the distortions caused by imperfect objective measures (Baker et al. [1994]). Levin [2003] modeled subjective performance measures assuming the principal can privately observe a noisy signal of the agent's performance which is not based on objective measures the agent can observe. When the principal subjectively evaluates the agent's performance, the usage of such measure may be abused - principal may understate the agent's effort and lower the transaction he should make to the agent.

To make the principal honestly evaluate the agent's performance, Levin [2003] proposed a system called full review contract under which the principal must provide a full performance evaluation at each period based on all existing historical information. That is, all private information is revealed after every period. Since both parties face the threat of terminating the relationship, the principal therefore is discouraged from under-evaluating the agent's performance. Levin [2003] proved that an optimal full review contract must make the same transaction at each period whilst permitting some possibility that the parties renege and break the contract. However, efficiency loss becomes inevitable even in the optimal contract.

Researchers have long been studying the use of subjective evaluation in incentive contract, although hardly can they find a better way to achieve an efficient outcome.[2] Some argue that internal mediation helps to resolve disputes caused by different beliefs in performance through narrowing the space of disagreement and helps to maintain relationships (Carver and Vondra [1994]). Baker et al. [1994] proposed a combine use of explicit contracting and implicit contracting to achieve positive payoff while neither of the two alone can yield positive payoff.

With subjective evaluation, MacLeod [2003] shows that the optimal contract entails the use of more compressed evaluations relative to the case with verifiable performance measures. By introducing a special mechanism to agents' such that they are able to conduct a costly punishment to the principal whenever they feels the principal's assessment is unfair, MacLeod [2003] shows the contracting cost can be further reduced.

Kambe [2006] studied a situation where, the principal has the option to choose objective measurement or subjective measurement, once the objective measurement is made, then the payment will not be influenced by the principal's subjective

[2]Unlike objective performance measure, parties may have different beliefs on subjective performance measurement, which causes distortion in players' behavior.

measurement. He shows that contracts involve subjective component only when subjective measurement is sufficiently accurate or only subjective measurement is available, despite the possibility that it may be abused.

Maestri [2012] compares bonus payment verses efficiency wage under subjective evaluation. In bonus payment contract, the agent receives a bonus after each period based on the subjective measure of performance made by the principal, the agent uses his private information to review the principal's reward, the agent quits the relationship once he feels his performance is undervalued. In efficiency wage contract, the agent receives a high fixed wage and faces the threat of being fired if the output is low. Maestri [2012] shows that bonus-payment contract always dominates efficiency-wage contract and bonus payment is asymptotically efficient which means it can converge to first best outcome.

Most of the results are derived under the typical framework with only one agent and the accepted wisdom in using subjective performance measure will actually lead to efficiency loss once parties disagree on the measure. However, there is a well known situation where subjective measurement is successfully applied in a situation with multiple agents - tournament. The principal fixes the total payment to the agents, and each agent receive a payment based on his rank among the team. Since the total payment is fixed, the principal has no incentive to under-evaluate the agents' performances.

Based on the existing literature of team moral hazard, Rayo [2007] extends Levin's model by introducing implicit relational incentive into a team setting with profit sharing and shows how implicit and explicit contract interact. Rayo [2007] studies two situations where effort is either perfectly observable or can only be observed as a noisy signal. Rayo [2007] shows that under the first situation, the optimal contract divides the output to all team members. On the other hand, if effort signal is sufficient noisy, profit sharing and relational incentive become perfect substitution with each other, the optimal contract thus concentrate all output into one of the team members who is more difficult to provide implicit incentive to, and all other team members receive only relational incentives. Although Rayo [2007] said the model is budget balancing, he introduced a special individual who does not exert effort. This individual plays the exact role of the principal in Holmstrom [1982], providing the possibility of breaking the budget when output is low.

When considering repeated agent-interaction with implicit contracting, joint performance evaluation can be utilized since it equips the agents with tools to punish their colleagues and reduce the principal's cost of providing incentives (Che and Yoo [2001]). Kvaloy and Olsen [2006] showed that the optimal incentive scheme is determined by the productivity of agents: the higher the productivity of agents the more frequently relative performance evaluation is used.

All the previous results are obtained under one important assumption that agents' strategies form a Nash equilibrium. However in repeated team work, the agents may coordinate their actions or may even form collusive strategies, especially in relative performance measure scheme. If the agents' payments are totally conditioned only on their ranking in the team, they can collude and jointly be better off while spending low effort. Principal's exogenous wage scheme thus have no use since the agents' form an endogenous wage allocation. Kvaloy and Olsen [2006] studied the problem of collusion caused by the relative performance payment scheme. Interestingly, the basic results remain the same. For sufficiently high productivity the relative payment scheme is still optimal even with an allowance for collusion. However, the possibility of collusion among agents increases the cost of relative payment scheme leaving more possibility for joint performance measure.

Although classical economic theory assumes that individuals are self-interested and maximize their own utility while having no consideration of other's well being, more and more empirical and experimental findings do not confirm this view. It has been widely found that people have fairness concerns and want to be treated fairly. Economists have long been studying how to incorporate these physiological findings into modern economic theory. Bolton and Ockenfels [2000], Fehr and Schmidt [1999] proposes inequity aversion models where inequality in payoffs causes disutility; people whose payoff is lower feel envy and whose payoff is higher become altruistic. Rabin [1993] studies an intentions-based model where people judge motivation from the outcome and tend to help those who treat them well and hurt those who treat them badly. Suggesting agents care more about bad outcomes caused by other players actions rather than the nature. There are other researches introducing social norms into fairness concerns (Huang and Wu [1994]).

As discussed above, inequity aversion plays an important role in people's decision making and should be brought into consideration in incentive provision. People with inequity aversion tend to take strategies that increase equity among groups or within society. Although strict equal allocation is rarely seen in the real world, people would prefer an equal allocation when total group profit becomes irrelevant in their utility functions (Engelmann and Strobel [2004]). Two kinds of games are often used in the study of inequity aversion, the dictator game where subjects have no incentive to transfer any amount of money and the ultimatum game where subjects have no incentive to reject money. When assuming inequity aversion, however, they will spontaneously transfer and reject money which is entirely away from the prediction of classical theory.

To build up contract theory with inequality aversion, it's essential to determine how can inequality aversion be represented in people's utility functions. Economists model inequity aversion in different ways, one is self-centered inequity aversion proposed by Fehr and Schmidt [1999]. The inequity aversion in their model comes from comparison with others. Such individual centered inequity aversion may be affected by the environment. If players can sanction other group members, even though only a minority of group members has inequity aversion, fair outcome can be achieved. Another well known model was presented by Bolton and Ockenfels [2000], quite similar to Fehr and Schmidt [1999]'s model, but in B Bolton and Ockenfels [2000]'s model the relative magnitude of parameters that determine the intensity of inequity aversion are influenced by the tension between profit maximization and social comparison. Rabin [2002]combined these two models into a two-person, three-parameter model which functions in a similar way of the previous two.

Among all the experiments exploring players' preferences towards inequality aversion, we are typically interested in the linear public good model. Since linear public good game can be viewed as a special case of team production where the output is equally shared among all agents. When building inequality aversion into the linear public good model, Fehr and Schmidt [1999] found that cooperation exists but with low contribution levels unless a punishment scheme is introduced. This provides a vivid example on how inequality aversion can be utilized to enforce good outcomes. However, the model made several strong assumptions we would like to relax in this dissertation. Thus we would further polish the theory model before we apply it to team incentive provision.

While there do exist literatures about team contracts with other regarding preferences. Itoh [2004] studies the inequity aversion between agents in standard team contract, inspiring us to extend our focus to inequity aversion in implicit contracting. However, the existing literature on this issue is restricted either to deterministic production technologies, binary effort decisions or the focus of their analysis is not on inequality aversion but envy, i.e. the worker cares only about being worse off and not about being better off. However, empirical and experimental evidences have confirmed the effect of altruism on human behaviors.

Inequity aversion brings distortion to incentive provision in one-period setting. In relational contract, inequity aversion affects incentive compatibility and individual rationality in the same way as in standard team moral hazard model, however, it may enable the principal to provide less incentive to make the contract self-enforceable. For agents with inequity aversion, reneging brings them larger utility loss compared with self-interested agents.

Chapter 3

Relational Contract in Team Production and Linear Sharing Rules

3.1 Introduction

This chapter deals with moral hazard problem in teams. Although existing literatures established abstract sharing rules that has been proved to be optimal, complicated sharing structure are hardly seen in reality. Most team and partnership contracts we observe are linear, which merely specifies proportional shares for each participant and some times constant transfer of payments. One possible explanation of the contradiction from theory to application is that complicated contracts are more difficult to be enforced. No matter how clear a contract is understood by the contractual parties, without enforcement from the legal system, it is powerless.

Built upon the previous works by Holmstrom [1982] Levin [2003] and Rayo [2007], this chapter tries to fill the gap of Rayo [2007]'s research. We first investigate the usage of relational contracting under deterministic production. We showed that in deterministic case, the noisy soft information does not play an important role in contracting, the final output is strong enough to deter possible deviation in effort level among agents and trigger strategy can be used by depending trigger on the final output. Since the self-enforcing property limit the level of punishment in relational contracts, we show that under certainty, trigger strategy based on output itself weakly dominate relational contract.

The latter half of the chapter follows Rayo [2007]'s model, while trying to relax the linearity and non-negative assumptions Rayo [2007] made to the model. We first justify the optimality of linear sharing rule when first-order approach is valid and negative shares are not allowed. We then move on to allow negative sharing, by discussing the validity of first-order approach with negative shares. We find the optimal sharing would have a special "franchising principal" who plays a very different role compared with Rayo [2007]'s result. However, the conditions asserted by Rayo [2007] is not sufficient enough to ensure the validity of first-order approach. Little is known about what conditions are needed to ensure first-order approach can be applied in a team case.

3.2 The Model

We consider a team with n members who take part in a joint production repeatedly. At the beginning of each period, each team member i simultaneously decide whether or not to participate in the production process. Let $d_i^t \in \{0, 1\}$ denote the participation decision for each team member i at each period t. For an individual i who is willing to take part in the production at t, we have $d_i^t = 1$ and $d_i^t = 0$ otherwise. For the production to take place at period t, we require that all the team member must involve, that is $\prod d_i^t = 1$.[1]

At the beginning of each period t, the agents sign a court-enforceable agreement, which specifies: the participation decisions for all agents[2] and how the final output will be shared among all the participating agents. Let S^t denotes the sharing rule[3]. Then each team member takes an unobservable action $a_i^t \in \mathbb{R}^+$ which incurs a cost $c_i(a_i^t)$ to the individual i. The cost is increasing, differentiable and convex in a_i^t. The action of all the team members $a^t = (a_1^t, a_2^t, \ldots, a_n^t) \in \mathbb{R}_n^+$ jointly determines the final output Y^t according to a production function $f : \mathbb{R}_n^+ \to \mathbb{R}^+$. The production function f is increasing, differentiable and concave in the team members' actions a^t.[4] At the same time, each team member's action produces a noisy signal l_i which is publicly observable but cannot be verified by a third party.

[1] This is not uncommon in partnerships where every member is crucial. However, in multi-agent models it may not be the case.

[2] If an agent decides to participate in the production in period t, then positive effort is expected from this agent and we assume there exist technology to tell whether an agent shirks completely (0 contribution).

[3] Here we would treat the generation of the sharing rule as an external procedure which will not be considered in the game. It can possibly suggested by one agent or several agents jointly, or it can be an outcome of group negotiation.

[4] Here we implicitly assume that the production function is stationary through out the time period.

Assume the signals following the same conditional density function $g(l_i|a_i)$ among all team members, with c.d.f $G(l_i|a_i)$. The outside option for each individual is denoted by $\bar{\pi}_i$.

At each period t, each team member observes the output Y^t but cannot observe others' actions. The production function, the sharing rule and each member's cost functions are common knowledge. Each team member i's pay-off under a sharing rule S^t is given by:

$$\pi_i^t = s_i^t(Y^t) - c_i(a_i^t) \tag{3.1}$$

The total surplus of the team is given by:

$$\Pi^t = Y^t - \sum_i c_i(a_i^t) \tag{3.2}$$

In each period, the greatest surplus the team will be able to achieve is: $\Pi^* = \max Y^t - \sum_i c_i(a_i^t)$, we denote this surplus as the first best surplus and the action vector $a^* = \arg\max\{Y^t - \sum_i c_i(a_i^t)\}$ denotes the first best action vector. Here we focus on Rayo [2007]'s routine to provide an extensional research under a wider range of situations, while the discussion of whether utility is still transferable with relational contract will be left to the next chapter[5]

[5] Rayo [2007] assumes the utility is transferable between agents and therefore he can focus on maximizing the total surplus without worrying about how the surplus need to be distributed among agents. However if agents' utilities are not transferable, there will exist some outcomes that are strictly preferred by some agents while not maximizing total surplus.

3.2.1 Linear Sharing Rule

Here we would restrict our attention to a special kind of sharing rule which is most commonly seem in real life - linear sharing rule.

Definition 3.1. A sharing rule $\mathcal{S} : \mathbb{R} \to \mathbb{R}^n$, where $n = 2, 3, \dots$ is said to be linear if:

1. For any $Y \in \mathbb{R}$, the share to each team member i has the following structure: $s_i(Y) = \alpha_i Y + \beta_i$, where α_i and β_i are constants or functions that are independent of Y, where the share is non-negative: $\alpha_i \geq 0$.
2. The budget is balanced such that $\sum \alpha_i = 1$ and $\sum \beta_i = 0$.

The definition says, a linear sharing rule contains two parts, for each team member i it specifies a portion α_i of the output which must sum up to 1 and a constant income/payment β_i which must sum up to 0. The simple structure of linear sharing rules makes them easily enforced by courts. Moreover, by using a linear sharing rule the team members are pretty sure that the shares they get $s_i(Y)$ is monotone increasing with the output Y, since $S_i'(Y) = \alpha_i \geq 0$.

However, the incentive provided by linear sharing rule is not strong enough to produce the first best outcome (Holmstrom [1982]). This is due to the fact that any agent's shirk will save his/her effort cost fully while the negative impact on the output will be shared with others. If we stick to balancing the budget, the punishment from a linear sharing rule can never be strong enough to achieve an efficient outcome. In the static model in Holmstrom [1982], the action vector of the team $\hat{a}$ constitutes a Nash equilibrium if and only if for each i, $\hat{a}_i$ solves $\max s_i(f(a_i, \hat{a}_{-i})) - c_i(a_i)$, subject to for each i, individual rationality is satisfied that is $\pi_i \geq \bar{\pi}_i$.

3.2.2 Infinite Repeated Game and Relational Contract

With the refinement of perfect equilibrium, we are pretty sure that the efficiency can be improved when a game is played infinitely (although first-best efficiency might still not be achieved). The threat of falling back to the non-corporative static equilibrium gives incentive to every team member to maintain a corporative outcome for each period. This idea of implicit cooperation is introduced to contracting which is referred to relational contract (Levin [2003]). In a relational contract, terms can be enforced implicitly other than explicitly (i.e. through a court), providing more flexibility of contracting. The long-term benefit from keeping to a productive relation can prevent players pursuing short-term interests through deviating. A relational contract contains the following elements:

1. A court-enforceable sharing rule. If the production occurs at period t, a formal contract specifying how the output will be shared should be established. If the sharing rule is linear, it then contains a share variable α^t and a fixed income/payment β^t.

2. A discretionary payment p. Since the payment is discretionary, p may depend both hard and soft information. Still, the budget within the team must be balanced such that $\sum p_i^t = 0$ for each t.

3. The effort level each team member should take.

4. An action d^t for all the team members specifying whether or not to join the production at period t.

5. An action φ^t for all the team members specifying whether or not to make the discretionary payment at period t.

6. The behavior after any of the team members renege. If one of the players renege, it's natural for his opponents minimax this member's pay-off, that is the trade been terminated and everyone receives their outside option, even though this punishment is costly (all the other team members will only have their outside options themselves afterwards).

Note that we are currently focusing on deterministic case where the actions each team member makes determine a value of joint output. For non-deterministic case, which will be discussed later, team members' actions determine a distribution rather than a value of output.

Let $H^t = \{h_1, h_2, \ldots, h_t\}$ denote the pubic history at beginning of period t, $h_t = \{Y^{t-1}, S^{t-1}, l_1^{t-1}, \ldots, l_n^{t-1}, p_1^{t-1}, \ldots, p_n^{t-1}, d_1^{t-1}, \ldots, d_n^{t-1}, \varphi_1^{t-1}, \ldots, \varphi_n^{t-1}\}$ denote the public information from the end of period $t-1$ to the beginning of period t, and define $h_1 = \varnothing$. We say a relational contract is self-enforcing if for any public history, they are willing to execute the discretionary payment p_i^t, that is $\varphi_1^t = \cdots = \varphi_n^t = 1$

Thus, a self-enforcing relational contract must satisfy the following constraints:

1. All team members must gain at least much as their outside options, that is for all i,

$$\pi_i \geq \bar{\pi}_i \tag{3.3}$$

2. The target effort level is implementable, that is IC constraints must hold. For all i and t,

$$a_i^t \in \arg\max E[S_i^t(y^t) + p_i^t(H^t, l^t, y^t) - c(a_i^{t\prime}) + \frac{\delta}{1-\delta}\pi_i(H^t, l^t, y^t)|H^t, a_i^{t\prime}, a_{-i}^t] \tag{3.4}$$

3. All team members are willing to implement the discretionary payment p_i^t in all period t:

$$p_i^t(H^t, l^t, y^t) + \frac{\delta}{1-\delta}\pi_i(H^t, l^t, y^t) \geq \frac{\delta}{1-\delta}\bar{\pi}_i \tag{3.5}$$

The individual rationality constraint 3.3 is a necessary condition for the production to initiate. Constraint 3.4 states the decision problem each team member faces when choosing the optimal effort level: the action they choose must not only maximize the current payoff which includes the share of output, the discretionary payment and the cost of effort but also the influence on future payoffs (as future payoffs can contingent on public history). Constraint 3.5 says that if a is to be sustained as equilibrium, then every team member must prefer to pay or receive the discretionary payment and remain the relationship (LHS) than to renege and receive outside option onwards (RHS).

Since the fixed payments do not influence the incentive compatible constraints, Levin [2003] states that we can first focus on maximizing the total surplus without considering the individual rationality constraints, the fixed payment can be used to redistribute the output to make sure every team member get at least as much as his/her outside option without affecting the incentive provision.

Lemma 3.2. *If an optimal relational contract exist, then there exists a stationary contract that is optimal.*[6]

Proof. . See Appendix A □

With lemma 3.2 we can now focus on stationary relational contract without loss of generality. An optimal stationary relational contract can be described as: For any period t the contract solves the maximization problem,

$$\max y^t - \sum c(a_i)$$

subject to the following constraints for each i:

$$a_i \in \arg\max E[s_i(y^t) + p_i(l^t, y^t)|a_i', a_{-i}] - c(a_i')$$

[6]The proof of the lemma to the non-deterministic case is similar, the proof would hold by substituting output y into expected output.

$$p_i(l^t, y^t) + \frac{\delta}{1-\delta}\pi_i(l^t, y^t) \geq \frac{\delta}{1-\delta}\bar{\pi}_i$$

Note that although the enforcement of a relational contract also depends on the game being played infinitely, the incentive provision has significant difference towards explicit contracts. In a stationary relational contract, incentive is paid through discretionary payment period by period. Namely, everything is solved in the current period and self-enforcing constraint makes everyone willing to carry out the discretionary payment to each other in fear of breaking the long-run benefit. While in an explicit contract, payment can only be contingent to verifiable output y. The folk theorem tells us that when a game is played infinitely, players will work cooperatively to a more efficient outcome in the fear of falling back to the worst static equilibrium. In team setting, Winter [2009] shows that once the production function is concave, increasing one agent's target effort level would not distort the incentive provision to other agents. In our model, each team member's incentive can be divided into two parts: the share of the output and the discretionary payment. Let the difference of payment δv denote as the power of implicit incentive.[7]

[7] Since for any given y and l, the discretionary payment must sum up to 0 across all agents. Thus we can generalize the expected discretionary payment on the equilibrium path as 0 and focusing on solving the payment function off the equilibrium path, as only the incentive power matters

3.3 Deterministic production: Trigger strategy

In deterministic case, although the output y doesn't tell us any information on who reneges, it is sufficient to deter any deviation on effort. Hence, y is enough for conduct punishment strategies in infinitely repeated game, we can construct the following trigger strategy based on the output y

Theorem 3.3. *With deterministic production, for sufficiently large discount factors, there exists a set of feasible sharing rules under which first-best outcome can be achieved by taking advantage of trigger strategy. The contract is characterized as follows. For all i and $t > 1$, the team members' strategies for whether or not to participate are*

$$d_i^t = \begin{cases} 1 & \text{if } y^{t-1} \geq y(a^*) \\ 0 & \text{if } y^{t-1} < y(a^*) \end{cases}$$

, while a^* is the first-best effort level.

Proof. See Appendix A □

The idea of terminate the relationship share some similarity of Holmstrom [1982]'s punishing the whole group. By assuming that the production cannot take place if any one of the team members does not participate, we would ensure that the enforcement is robust to any side-contract or collusion. Note that the theorem also holds even if we relax the assumption that the production needs involvement from all the n members. Even termination is costly to other individuals who do not shirk, they will still stick to the strategy to enforce the best outcome.

We also notice that both trigger strategy and relational contract's incentive provision comes from the value of continuation. We would formally compare these two in the following theorem.

Theorem 3.4. *With deterministic production, if* $\mathbf{a}^*$ *can be implemented by a relational contract, it can also be implemented by trigger strategy.*

Proof. See Appendix A □

Theorem (3.4) shows that under certainty, relational contract is weakly dominated by trigger strategy. As in deterministic case, the output level becomes a perfect signal on reneging on effort level of the whole group and trigger strategy provides the harshest punishment to the whole group compared with the relational contract. Thus in deterministic case, the design of relational contract would not be very useful. However, when we continue to investigate under the uncertainty of production, things can be different. The most significant advantage of relational contract under uncertainty is: production can always be continued. While for trigger strategy, no matter where the threshold of trigger is set, there are always possibilities of stopping the production even if no one reneges. In the following section we will continue Rayo [2007]'s discussion on the relational contract under uncertainty, but focus would be on the sharing rule.

In general, trigger strategy provides a sample of how severe the punishment of deviation can be in the most extreme case. Conducting trigger strategy automatically ends the relationship and thus binds the self-enforcing constraint. However we do not like to see relationship end due to pure stochastic factors that are not controllable by agents. In this sense we may use trigger strategy in deterministic setting, but in certain circumstances, relational contract would be more preferable.

3.4 Non-deterministic production: Sharing rule in relational contract

In this section we assume the output is randomly determined by the effort level $\mathbf{a}$, all the properties of the production function in the previous section now applies to the expected output function $E[y|\mathbf{a}]$. With uncertainty, the output level would no longer be sufficient enough to tell whether any agent shirk on the effort, thus the individual signal can be valuable and we shall discuss the design of the relational contract. As Rayo [2007] has discussed the optimal implicit incentive given linear sharing, we would continue our discussion on the optimal explicit incentive given the following assumptions hold:

Assumption 3.1. *The performance signal l is a sufficient statistic for a relative to (y, l)*

This assumption ensures that the discretionary payment can be expressed as a function of l without loss of generality.

Assumption 3.2. *The discretionary payment $p_i(\mathbf{l})$ is additively separable: $p_i(\mathbf{l}) = w_i(l_i) + u_i(l_{-i})$*

As l_{-i} are independent on a_i, for every agent i, the term $u_i(l_{-i})$ in IC constraint can be treated as a constant. Therefore we can simplify our IC constraints to the following:

For each i

$$a_i \in \arg\max E[s_i(y)|a_i', a_{-i}] + E[w_i(l_i)|a_i'] - c(a_i')$$

Assumption 3.3. *The first-order approach is valid for any non-negative sharing rule.*

We take this setting from Rayo [2007], however we shall point out that the conditions provided by Rayo [2007] are not sufficient to show the validity of first-order approach. Rayo [2007] asserted the following two conditions in his paper:

Condition 3.1. *For all values of a_i, the likelihood ratio $\frac{g_{a_i}}{g}(l_i|a_i)$ is increasing in l_i*

Introduced by Mirrelees [1976], assumption 3.1 is also called the Monotone Likelihood Ratio Condition (MLRC). This assumption ensures that a_i has a monotonic impact over l_i, in terms of first-order stochastic dominance. Thus the discretionary payment w_i will be increasing with l_i.

Condition 3.2. *$G(l_i|a_i = c^{-1}(z))$ is convex in z for all l_i*

This assumption is a relaxed version of Rogerson [1985]'s convexity density function, which ensures the concavity of term $E[w_i(l_i)|a_i'] - c(a_i')$ in IC constraint.

Intuitively the density function $G(l_i)$ is decreasing in a_i, as the more effort agent exerts, the lower the probability that the signal l_i will be less or equal to a certain level. The assumption guarantees that the marginal impact of effort on signal, after netting the cost, will be decreasing.

However, the sufficiency can only be proved under single agent setting, namely when only one IC constraint need to be considered. Little has been studied of which conditions are needed for the validity of first order approach under team and multi-agent setting (Kim and Wang [1998]).

Assuming first-oder approach, we can replace the IC constraints with their FOCs. We would then show the minimum implicit incentive needed for target effort $\mathbf{a}^*$ given sharing rule $\mathbf{s}$.

Lemma 3.5. *Given target effort level* $\mathbf{a}^*$ *and sharing rule* $\mathbf{s}$*, if first-order approach is valid, the optimal* w_i *takes the following form:*

$$w_i(l_i) = \begin{cases} v_i, & \textit{for all } l_i \textit{ such that } g_{a_i}(l_i|a_i) > 0 \\ 0, & \textit{Other wise} \end{cases} \tag{3.6}$$

, where

$$v_i = \frac{c'(a_i^*) - \frac{\partial E[s(y)|a^*]}{\partial a_i}}{\int \max\{0, g_{a_i}(l_i|a^*)\} dl_i}$$

Proof. We first show that the one-step payment functional form stated in equation (3.6) is optimal.

When assessing agent i's effort level given signal l_i, we construct the following hypothesis test:

- H_0: $a_i \geq a_i^*$
- H_1: $a_i < a_i^*$

By the optimal Neyman-Pearson detection rule, we reject the null if and only if $g_{a_i}(l_i|a_i^*) \leq 0$. Thus it's optimal to punish agent i when $g_{a_i}(l_i|a_i^*) \leq 0$.

We then need to find the value of v_i. From the first-order condition of IC constraints, we have:

$$\frac{\partial E[s_i(y)|a^*]}{\partial a_i} + v_i \int \max\{0, g_{a_i}(l_i|a^*)\}dl_i =' (a_i^*) \tag{3.7}$$

Rearranging the equation we have $v_i = \frac{c'(a_i^*) - \frac{\partial E[s(y)|a^*]}{\partial a_i}}{\int \max\{0, g_{a_i}(l_i|a^*)\}dl_i}$ □

Lemma 3.5 is a general case of Rayo [2007]'s Lemma 3 without assuming linearity in the sharing rule. At this stage we only focus on the optimal discretionary payment function w_i with effort level and sharing rule given. The discussion of the self-enforcing constraint will be left after the revelation of optimal sharing rule.

Theorem 3.6. *If an optimal relational contract is described by* $\mathcal{S} = (s_1(.), \ldots, s_n(.))$ *and* $u(.)$ *targeting effort level* $\mathbf{a}^*$*, there exist a linear sharing rule such that for each* i $s_i' = \alpha_i y + \beta_i$*, where* $\alpha_i = \frac{\partial E[s_i(y)|a^*]}{\partial a_i} / \frac{\partial E[y|a^*]}{\partial a_i}$ *and* $\beta_i = E[s_i(y)|a^*] - \alpha_i E[y|a^*]$ *that is also optimal*

Proof. See Appendix A □

By proving the optimality of linear sharing rules, we justify the linearity assumption Rayo [2007] has made to his model. The optimality of linear sharing depends on the validity of first-order approach, however it's not clear what sufficient condition is needed to ensure the assumption.

We would now focus on the negative sharing extension to Rayo [2007]'s model. Let $\tau_i = P(u_i(l_i) \neq 0|\mathbf{a}^*)$ denote the probability that agent i's signal will satisfy condition $g_{a_i}(l_i|a_i) > 0$. Term $\int \max\{0, g_{a_i}(l_i|a^*)\}dl_i$ can be written as $\frac{d\tau_i}{da_i}$

Fixing the amount of incentive provided to each agent, it's shown by Rayo [2007] that the marginal rate of substitution between explicit incentive and implicit incentive is a constant $\frac{\partial E[y|a^*]}{\partial a_i}/\frac{d\tau_i}{da_i}$. While the amount of explicit incentive is limited by the budget balancing constraint, the optimization sharing rule should minimize the implicit incentive needed due to the self-enforcing constraint. When shares are non-negative, it's trivial that shares will be concentrate to one agent who is the most difficult to be provided with implicit incentive (Rayo [2007]) and all other agents receives no explicit incentive.

However, this sharing structure would collapse as soon as we allow negative sharing. Once agent's shares are no-longer bounded by 0 below, efficiency can be easily improved by reducing shares for agents who are easier to be provided with implicit incentive. Yet there is a special case under which the optimal sharing can still be positive even when we relax the non-negative sharing constraint.

Lemma 3.7. *Given α and $\mathbf{a}^*$, if for each agent i, $\frac{\partial E[y|a^*]}{\partial a_i}/\frac{d\tau_i}{da_i}$ is the same, then any arbitrary non-negative linear sharing rule is optimal.*

Proof. See Appendix A □

Lemma 3.7 provides a condition under which the design of sharing rule is not important as the self-enforcing constraint will not be affected by the choice of the explicit sharing rule. Therefore, under this special case whether allowing negative sharing does not matter.

However, for other cases when the sharing rule does matter and very likely negative sharing is needed to improve efficiency, we need to answer one question before searching for the optimal sharing rule. That is: Is first order approach is still valid? If not, then linear sharing may no longer be optimal. Dropping the non-negative sharing assumption doesn't mean the sharing is unbounded, rather it's having a wider range such that for all i, $\alpha_i \in [2-n, 1]$. The lower bound suggests that the minimum sharing an agent can get is $2-n$, when all other agents receive 100% share of the output.

Given α and $\mathbf{a}^*$, the second order derivative for agents' payoff function with respect to their effort level a_i is stated as below:

$$h_i = \alpha_i (E[y|\mathbf{a}^*])'' + v_i \tau'' - c''(a_i) \tag{3.8}$$

Lemma 3.8. *The marginal impact of α_i on h_i is monotone.*

Proof. See Appendix A □

Since the marginal impact of share on the second derivate of payoff function is monotone, if this impact is non-negative, then the second order condition will always satisfy even with negative shares. This requires the highly concavity of τ. On the other hand if $\frac{\partial h_i}{\partial \alpha_i} < 0$, due to the monotonicity, for each agent i there exist a $[\alpha_i]_{min}$ such that first order approach is still valid for $\alpha_i > [\alpha_i]_{min}$. For a special case if $[\alpha_i]_{min} < 2 - n$ for all i, first order approach will still be valid for the optimization problem.

Depending on the value of $\frac{\partial h_i}{\partial \alpha_i}$, α_i will be bounded below either by $2 - n$ or $[\alpha_i]_{min}$, we would then focus on the case when the validity of first-order approach is guaranteed, the optimal sharing structure can be extended to the other case applying different lower bound.

Theorem 3.9. *When first-order approach is valid and negative sharing is allowed, the optimal sharing would have the following property:*

- *For agent* $i \in \arg\min \frac{\partial E[y|a^*]}{\partial a_i} / \frac{d\tau_i}{da_i}$, $\alpha_i = 2 - n$
- *For any other agent* $j \notin \arg\max \frac{\partial E[y|a^*]}{\partial a_i} / \frac{d\tau_i}{da_i}$, $\alpha_j = 1$

Proof. See Appendix A □

Theorem 3.9 provides a completely contradictory solution to Rayo [2007] where sharing can only be non-negative. However, they do share a similarity such that the optimal sharing would make agents' shares binding to the two extremes. According to Rayo [2007], when negative sharing is not allowed, there is one agent who is the most difficult to be provided with the implicit incentive gets the full share of the output while all other agents' shares automatically binds to 0 due to the budget balancing constraint. However, when negative sharing is allowed, theorem 3.9

shows that more than one agent can get a full share, and there exist one (maybe more) agent who is the easiest to be provided with implicit incentive, concentrate all the negative shares and will be fully incentivized by relational contract. As comparing with this agent, it would be costly to provide implicit incentive to other agents and intuitively they will be receiving explicitly incentives only and negative sharing allows every other agent been fully motivated.

There are two key ingredients deciding who could be receiving a negative share: the marginal productivity of effort and the quality of the performance signal. The higher the marginal productivity the more effective explicit incentive would be. While the more volatile the performance signal, the more difficult it would be to provide the agent with relational incentive. Thus the agent who receive negative share must be less marginally productive while has a stronger and less noisy effort signal. To further illustrate our result consider the following examples

Example 3.1. *Consider a team consists of agents 1, 2 and 3. Let* $\mathbf{a} = \{a_1, a_2, a_3\}$*, with the output* $\mathcal{F} = 2a_1^{\frac{1}{2}}a_3^{\frac{1}{2}} + 2a_2^{\frac{1}{2}}a_3^{\frac{1}{2}}$*. Each agent's cost function is defined as follow:* $c_1 = a_1^2/2$*,* $c_2 = a_2^2/2$*,* $c_3 = a_3^2$*. Suppose the effort signal* l_i *follows the same conditional density function across all* i *. If we target the first best effort level* $\mathbf{a}^* = \{1, 1, 1\}$*. The marginal impact of each agent's effort on the output will be* 1*,* 1 *and* 2 *respectively. Given the fact that these agent's effort signal are equally noisy and agent* 3*'s effort has larger marginal impact on output level than the other two agents, if we decrease agent* 3*'s share, it would require a higher amount of relational incentive to compensate for the loss in incentive power. According to theorem 3.9, minimizing the amount of relational incentive would make agent* 3

get a full share of the output, while the other agents would get the rest. By theorem 3.7, the marginal rate of substitution of explicit and implicit incentive is the same across agent 1 and 2, how they share the output really doesn't affect the result as long as their shares sum up to 0 such that $\alpha_1 + \alpha_2 = 0$*. In this typical case, if the negative sharing is not allowed we would have a unique optimal sharing with* $\alpha_1 = \alpha_2 = 0$ *and* $\alpha_3 = 1$*, which is also exactly what's suggested by Rayo [2007]. If we allow negative sharing then we could let agent* 1 *get the negative share while other agents get the full share, with* $\alpha_1 = -1$ *and* $\alpha_2 = \alpha_3 = 0$*. The only difference between our result with Rayo's is that in Rayo [2007]'s case only one agent is not receiving implicit incentive while in our case, there will only be one agent receiving implicit incentive.*

Theorem 3.9 also shed light on the possibility of monitoring, if there is a technique to accurately measure an agent's effort[8], then it would be worthwhile to conduct monitoring as this is obviously possible for one agent. From organizational structure perspective, whether this agent is or not contributing to the project doesn't matter, teams can always set up a job that the effort inputting can be easily measured. This typical agent act like a "team franchiser", the only difference is the payment to franchiser is made discretionally and after the output has been generated. In a way this special agent is more like a principal comparing to Rayo [2007], as the effort can be provision of asset while no actual labor effort is needed from this agent. For example the special agent provides a "recipe" of this typical food and a team is formed to formally produced and sell the food to the market. Theorem 3.9 shows that with relational contracting, the solution to moral hazard

[8]Or make the effort signal less noisy.

problem in teams only need the revelation of one agent's effort level, and this agent act like a real principal in terms of the nature of the work it's been specified endogenously by the model.

3.5 Conclusion

This chapter discusses the utilization of relational contract in both deterministic and non deterministic production. We showed that in the deterministic case the output level becomes a perfect signal on detecting reneging within the group, that the trigger strategy, which provides a means to punish the whole group on the continuation payoff, would weakly dominate the relational contract, which utilize the individual effort signal. While in non-deterministic case, we follow the discussion of Rayo [2007]. Based on Rayo [2007]'s result on the optimal relational incentive, we showed that the linear sharing rule is optimal. Furthermore, we discuss the optimal share agents would receive. When allowing for negative sharing, agents who are the easiest to be provided with implicate incentive will be motivated by implicit incentive only and usually receiving a negative share. The other agents will not receive implicit incentive and will receive a full share each.

Chapter 4

Relational Contract with Sub-teams

4.1 Introduction

In the previous chapter, we discuss the impact of stochastic individual performance signal in relational contracting. In large institutions however, it might be costly even just to make everyone's performance signal to be public information. This chapter will thus investigate an alternative setting, where we no longer focus on getting extra information from individuals. Rather, we would group the agents into sub-teams based on the tasks they are assigned to. Sub team performance then becomes the soft information we can utilize in relational contracting. Restricting the dimension of signal vector would also make deterministic signaling possible.

The emphasis of the chapter will be put upon the sub-team structure and the properties of the optimal sharing rule.

When modeling sub-team structure, we follow the deterministic model of Nandeibam [2002]. Here we extend the literature by introducing repeated agency with relational contracts to the model. Also, we formally define the classification mechanism between agents and sub-teams to further study the impact of the organizational structure on relational contracting.

We show that to allow precise punishment been fallen to the agent who shirks (any agent who deviate from the target effort level will definitely be punished implicitly) without breaking the budget, we can follow a simply criteria to form the sub-teams. The only requirement is: there must exist at least some heterogeneity between the sub-teams on organizational structure. Without being limited by the production technology, this allows us to implement a straightforward mechanism to detect who can result in a deviation without bringing the suspicion to the whole team (in which case it's unlikely to ensure the agent who shirks will be punished while keep budget balanced).

With the relational incentive been properly defined with the optimal grouping mechanism, we then focus our attention to the explicit incentive. We first analyze the necessary and sufficient condition for an outcome to be implemented with general sharing. We show that when each agent's share can be designed pointwise, utility will be transferable and we can always maximize the team's total surplus first and then distribute in any desirable way within the team.

We further compare the implementation under linear sharing with general sharing. As a key assumption adopted by Rayo [2007], linear sharing is widely seen in real world due to its simplicity. However the practicality of linear sharing is at the cost of freedom of distributing the surplus. We show that, with linear sharing agents' utilities are no longer transferable. Implying the optimal sharing depends on the social welfare function of the team. If the team only cares about the total surplus, then linear sharing would work just fine. On the other hand, we find that linear sharing would not be able to implement the efficient outcome specified by general sharing when the surplus is distributed very unevenly, i.e. principal-agent framework. All our result will be further visualized by a numerical example.

4.2 The model

We consider a production team that consists n agents, let $N = \{1, 2, \ldots, n\}$ be the set of agent. Each agent $i \in N$ can choose an effort level $a_i \in \mathbb{R}_{++}$[1]. We first investigate a deterministic production technology based on the joint effort level of all agents $\mathbf{a} = (a_1, \ldots, a_n) \in \mathbb{R}^n_+$.

On top of the team structure, the team members are further grouped into m sub-teams. Denote set $S = \{1, 2, \ldots, m\}$ as the sub-team set, we define the grouping mechanism as follows:

[1]Here we rule out the possibility of agents spending 0 effort because it's not a very interesting case. If an agent's optimal strategy is to spend no effort, it might be either due to low marginal productivity of the agent or low incentive payment within the team. It can be a two-sided screening problem in the team formation stage and will not be discussed here.

Definition 4.1. A grouping mechanism is a function $\mathbf{G} : S \rightarrow \mathcal{P}(N)/\emptyset$, where $\mathcal{P}(N)$ is the power set of N, with $\bigcup_{j=1}^{m} G(j) = N$

It's possible that two different sub-teams been mapped to the same subset of N. This counts for the fact that, although a same group of people works together, they might produce multi-dimensional signals. For simplicity we would separate them and treat them as different sub-teams to ensure the sub-team performance to be single dimensional.

The grouping information $\mathbf{G}$ is a soft information that is only known within the team but cannot be verified by outsiders. After being attributed to their sub-teams, each team member i can take an action a_i from his/her action space A_i, which incurs a cost $c_i(a_i)$ to the agent. The team members' actions $\mathbf{a} = (a_1, a_2, \ldots, a_n) \in A$ jointly determined each sub-team's performance. For each sub-team j, its output is determined according to function $f_j : A \rightarrow \mathbb{R}^+$, with the following property: $\frac{\partial f_j}{\partial a_i} > 0$, $\frac{\partial f_j^2}{\partial^2 a_i} < 0$ for all $i \in \mathbf{G}(j)$; $\frac{\partial f_j}{\partial a_i} = 0$ for all $i \notin \mathbf{G}(j)$.

Namely, each team member can only influence the performance of the sub-team(s) they belong to. The output of the team $\mathcal{F}$ is strictly concave for all values of $\mathbf{a}$. At the end, each member gets a share of the output $\mathcal{F}$ as his/her payoff, according to a pre-determined sharing rule $\mathcal{S}$. For any agent i, let $s_i(\mathcal{F})$ denote the monetary payoff he can get from output $\mathcal{F}$. Within the team, budget must balance, that is, $\sum_{i=1}^{n} s_i(\mathcal{F}) = \mathcal{F}$.

Without considering any enforcing problem, the sharing rule can depend on any public information. We say a sharing rule $\mathcal{S}$ sustains an effort vector $\mathbf{a}$ if for all

i, $\mathbf{a}_i \in \arg\max s_i(\mathcal{F}, f_1, \ldots, f_m) - c_i(a_i)$. Each agent i has an outside option $\bar{\pi}_i$ which represent his/her payoff if not participate in the production. The team's problem is to design a proper sharing rule to maximize the aggregate surplus within the team:

$$\max_{\mathcal{S},\mathbf{a}} \mathcal{F}(\mathbf{a}) - \sum_{i=1}^{n} c_i(a_i)$$

Subject to:

For each i

$$s_i(\mathcal{F}, f_1, \ldots, f_m) - -c_i(a_i) \geq \bar{\pi}_i$$

$$\mathbf{a}_i \in \arg\max s_i(\mathcal{F}, f_1, \ldots, f_m) - -c_i(a_i)$$

We denote the solution of the above problem $(\mathcal{F}^s, \mathbf{a}^s)$ as the static optimal solution, $\mathbb{S}$ as the set of all balanced sharing rules that sustains a^s as Nash equilibrium.

4.3 Relational contract

Nandeibam [2003] has shown how efficient outcome can be achieved by utilizing sub-teams' performances together with the grouping information **G**. In most cases, although the team members are aware of the sub-teams' outputs, it's difficult for an outsider to verify them. When sub-team outputs become soft information, all sharing rules in S that depend on sub-team performance will no longer be implementable in static situation. Thus, we introduce the concept of relational contracting to cope with the enforcement problem.

We define our relational contract the same way as the previous chapter, however our information set would look slightly different. Let d denote the action for all the team members specifying whether or not to join the production, τ be the action for all the team members specifying whether or not to make the discretionary payment, $H^t = \{h_1, h_2, \ldots, h_t\}$ denote the history information within the team at beginning of period t,

$$h_t = \{\mathcal{F}^{t-1}, S^{t-1}, f_1^{t-1}, \ldots, f_m^{t-1}, p_1^{t-1}, \ldots, p_n^{t-1}, \\ d_1^{t-1}, \ldots, d_n^{t-1}, \varphi_1^{t-1}, \ldots, \varphi_n^{t-1}\}$$

denote the public information from the end of period $t-1$ to the beginning of period t, and define $h_1 = \varnothing$. We say a relational contract is self-enforcing if for any public history, they are willing to execute the discretionary payment p_i^t, that is $\varphi_1^t = \cdots = \varphi_n^t = 1$.

In a dynamic game the team is maximizing it's discounted surplus stream[2]:

$$\max_{\mathcal{S}(t),\mathbf{p},\mathbf{a}} (1-\delta) \sum_{t=1} \delta^{t-1} \{\mathcal{F}(\mathbf{a}^t) - \sum_{i=1}^{n} c(a_i^t)\}$$

Subject to:

[2] Here I multiply the surplus by $1-\delta$ to take the per-period average surplus so that this can be compared with the surplus in the static game.

For each i, t and H^t

$$s_i^t(\mathcal{F}(a_i^t, \mathbf{a}_{-i}) + p_i^t - c(a_i^t) \geq \bar{\pi}_i$$

$$\mathbf{a}_i^t \in \arg\max s_i^t(\mathcal{F}(a_i^t, \mathbf{a}_{-i}) + p_i^t - c_i(a_i) + \frac{\delta}{1-\delta}\pi_i(H^t, f_1^t, \ldots, f_m^t)$$

$$p_i^t + \frac{\delta}{1-\delta}\pi_i(H^t, f_1^t, \ldots, f_m^t) \geq \frac{\delta}{1-\delta}\bar{\pi}_i$$

Lemma 4.2. *If the optimal contract exist, there is a stationary contract that is optimal*

Proof. See Appendix B □

With stationarity, the problem is reduced to:

$$\max_{\mathcal{S},\mathbf{p},\mathbf{a}} \mathcal{F} - \sum_{i=1}^{n} c_i(a_i)$$

Subject to:

For each i,

$$s_i(\mathcal{F}(a_i, \mathbf{a}_{-i})) + p_i(f_1, \ldots, f_m) - c_i(a_i) \geq \bar{\pi}_i \tag{4.1}$$

$$\mathbf{a}_i \in \arg\max s_i(\mathcal{F}(a_i, \mathbf{a}_{-i})) + p_i(f_1, \ldots, f_m) - c_i(a_i) \tag{4.2}$$

For any values of $(f_1, \ldots, f_m)$,

$$p_i(f_1, \ldots, f_m) + \frac{\delta}{1-\delta}\pi_i \geq \frac{\delta}{1-\delta}\bar{\pi}_i \tag{4.3}$$

Note that the self-enforcing constraint (4.3) has to hold for every possible values of sub-team performance. As the IC constraint relies on the fact that any punishment due to striking on effort can be enforced, if self-enforcement is merely defined on the performance on the equilibrium path, the relational contract would not work.

As a mechanism to redistribute team surplus on top of the sharing rule, the implicit transfers can be designed from two aspects; first, given the production information, who should we punish, who should we reward? Second, how severer(high) should we punish(reward) them?

The first aspect is directly related to the information structure, or more precisely speaking, the grouping mechanism $\mathbf{G}$. The key issue is, if an agent shirks, can we make sure that this agent is punished. As suggested by Holmstrom [1982], the easiest thing to do is to apply the punishment to the whole team by introducing an outside budget breaker. If we are to keep the budget balanced within the team, other mechanisms are required.

Definition 4.3. Given a group mechanism $\mathbf{G}$, a monitoring scheme that aims at detecting agent's deviation in effort can be described as a function: $\mathfrak{h}^{\mathbf{G}} : \mathbb{R}^m \to \mathcal{P}(N)$. We say a monitoring scheme fails if for some values of $(f_1, \ldots, f_m)$, $\mathfrak{h}^{\mathbf{G}}(f_1, \ldots, f_m) = N$.

For all target effort level $\mathbf{a}^*$ and any agent i, we say a monitoring scheme is perfect if for any $a_i' \neq a_i^*$, $i \in \mathfrak{h}^{\mathbf{G}}(f_i(a_i', a_{-i}^*), \ldots, f_m(f_i(a_i', a_{-i}^*)) \neq N$

Monitoring scheme is basically a mechanism to detect "who could be the one" when shirking occurs (based on sub-team performances). The condition for perfect monitoring implies that any agent's deviation given his/her peer effort level will make him/her been suspected while it's impossible for any agent to make the monitoring information totally useless.

Theorem 4.4. *There exist a perfect monitoring scheme if* $\bigcap_{j=1}^{m} G(j) \neq N$.

Proof. Given the target effort level $\mathbf{a}^*$ and suppose the grouping mechanism satisfy property: $\bigcap_{j=1}^{m} G(j) \neq N$, we will create a monitoring scheme that is perfect.

we first classify the agents into two groups A and B: for any agent $i \in N$, if $\frac{\partial f_j}{\partial a_i} = 0$ for at least one $j \in M$ then $i \in A$; if $\frac{\partial f_j}{\partial a_i} > 0$ for all $j \in M$ then $i \in B$, such that agents are classified based on whether they have influence over all sub-teams.

The monitoring scheme is defined as follows:

$$\mathfrak{h}^{\mathbf{G}}(f_1, \ldots, f_m) = \begin{cases} \emptyset & \text{if for all } j \in M\text{:}f_j = f_j(\mathbf{a}^*), \\ B & \text{if for all } j \in M\text{:}f_j \neq f_j(\mathbf{a}^*), \\ \bigcap_{j \in M'} G(j) / \bigcup_{k \notin M'} G(k) & \text{if for some } j \in M\text{:}f_j \neq f_j(\mathbf{a}^*) \end{cases} \tag{4.4}$$

, where $M' = \{j | f_j' \neq f_j(\mathbf{a}^*)\}$

With this monitoring scheme, for any agent i, if $i \in B$, a deviation from a_i^* would result in $f'_j \neq f_j(\mathbf{a}^*)$ for all $j \in M$, so that $i \in \mathfrak{h}^{\mathbf{G}} = B$. Given $\bigcap_{j=1}^{m} G(j) \neq N$, it's easy to verify that $B \neq N$.

If $i \in A$, a deviation from a_i^* will result in $f'_j \neq f_j(\mathbf{a}^*)$ where $i \in G(j)$ while $f'_k = f_k(\mathbf{a}^*)$ for all k that $i \notin G(k)$. As $i \in \bigcap_{j \in M'} G(j)$ while $i \notin \bigcup_{k \notin M'} G(k)$, we have $i \in \mathfrak{h}^{\mathbf{G}}$. Since $\bigcup_{k \notin M'} G(k) \neq \emptyset$, $\mathfrak{h}^{\mathbf{G}} \neq N$.

□

Theorem (4.4) provides a weak but practical condition to ensure the existence of perfect monitoring scheme under deterministic case, that is, there are at least some structural difference among the sub-teams. In extreme case where each sub-team contains the whole team members, the sub-team structure is collapsed into a normal team framework with multi-dimensional output. However our focus is trying to get the best outcome possible and have solved for the "optimal" sub-team structure. From now on, we will assume G satisfies $\bigcap_{j=1}^{m} G(j) \neq N$ and let $\mathfrak{h}^*$ denote an arbitrary perfect monitoring scheme.

Example 4.1. *To see how the monitoring scheme (4.4) works, consider a simple team with 3 agents 1, 2 and 3 together with 3 sub-teams* **A**, **B** *and* **C** *where the grouping of the sub-teams are described as follows:* $\mathbf{G}(\mathbf{A}) = \{1, 2, 3\}$, $\mathbf{G}(\mathbf{B}) = \{1, 2, 3\}$, $\mathbf{G}(\mathbf{C}) = \{2, 3\}$, *notice the condition in theorem (4.4) is met. And we can check any individual's shirking on effort will lead him been punished and at least one other team-member can balance the budget.*

If agent 1 shirk, it would lead to sub team **A** *and* **B***'s performance different from the target. According to the third case in the monitoring scheme, we first take the union of the team-members of sub team* **A** *and* **B** *to get all the possible agents who could result in the outcome which are agents 1, 2, 3 together. Then we exclude the agent that's impossible to shirk, from the fact that sub team* **C***'s performance is as expected, agent 2 and 3 cannot shirk, in this case agent* 1 *will be punished.*

If agent 2 or 3 shirks, then all the sub teams' performance would be different from target, from the monitoring scheme then punishment will only be conducted on agents who have influence on all sub teams. If either of agent 2 or 3 shirks, the monitoring scheme would punish both agent and let agent 1 balancing the budget.

With monitoring scheme, we restrict the number of agent we shall punish. The question then is to decide how we should punish them under the self enforcing constraint (4.3). Our next objective is to find the optimal level of punishment by maximizing the incentive power while ensuring the contract is self-enforcing.

With the perfect monitoring scheme, we will demonstrate the derivation of the optimal implicit incentive in the following theorem:

Theorem 4.5. *For any sharing rule S and target effort vector $\mathbf{a}^*$, for any agent i the following discretionary payment function*

$$p_i = \begin{cases} 0 & \text{if } \mathfrak{h}^{\mathbf{G}}(f_1, \ldots, f_m) = \emptyset, \\ \frac{\delta}{1-\delta}(\bar{\pi}_i - \pi_i) & \text{if } i \in \mathfrak{h}^{\mathbf{G}}(f_1, \ldots, f_m), \\ w_i(\mathfrak{h}) & \text{if } i \notin \mathfrak{h}^{\mathbf{G}}(f_1, \ldots, f_m) \neq \emptyset \end{cases} \tag{4.5}$$

where $w(\mathfrak{h}) \geq 0$ *will be optimal.*

Proof. See Appendix B □

Since the perfect monitoring scheme ensures the punishment of shirking will indeed fall on the agent who shirks, and budget can always be balanced, we are able to create the strongest relational incentive to the agents. The relational incentive can be specified as follows, if the final output equals the target output level, everyone receives 0 transfers. If someone shirks, then all agents been selected by the monitoring scheme will be fully punished, namely, binding the self-enforcing constraint where each agent $i \in \mathfrak{h}$ will get a negative transfer equal to $\frac{\delta}{1-\delta}(\bar{\pi}_i - \pi_i)$, while for all the other agent $j \notin \mathfrak{h}$ they won't get punished and will merely act as budget balancers with $p_j \geq 0$.

4.4 Efficiency under general sharing rule

The previous section discussed the provision of relational incentives. We showed that with perfect monitoring scheme, the implicit incentive can take a simple one-step functional form. The discretionary payments that contingent on sub-team performances will punish all agents who might shirk to the maximum. Given the target effort level $\mathbf{a}^*$, the sum of punishment is fixed: $\frac{\delta}{1-\delta}[\mathcal{F}^* - \sum_{i=1}^{n} c_i(a_i^*) - \sum_{i=1}^{n} c_i(a_i^*)\bar{\pi}_i]$. In this section, we will try to discuss the efficiency with general sharing rules.

With the clear specification of the relational incentive, the team problem becomes:

$$\max_{\mathcal{S},\mathbf{a}} \mathcal{F} - \sum_{i=1}^{n} c_i(a_i)$$

Subject to:

For each i and any $a'_i \neq a_i$

$$s_i(\mathcal{F}(a_i, \mathbf{a}_{-i})) - c_i(a_i) \geq s_i(\mathcal{F}(a'_i, \mathbf{a}_{-i})) - c_i(a'_i) + \frac{\delta}{1-\delta}(\bar{\pi}_i - \pi_i) \quad (4.6)$$

, where $\pi_i = s_i(\mathcal{F}(a_i, \mathbf{a}_{-i})) - c_i(a_i)$, rearrange the equation we have the new IC constraint:

$$\frac{1}{1-\delta}[s_i(\mathcal{F}(a_i, \mathbf{a}_{-i})) - c_i(a_i)] \geq \sup[s_i(\mathcal{F}(a'_i, \mathbf{a}_{-i})) - c_i(a'_i)] + \frac{\delta}{1-\delta}\bar{\pi}_i \quad (4.7)$$

Fix the payment each agent i receives at $\mathcal{F}^*$ as q_i, where $\sum_{i=1}^{n} q_i = \mathcal{F}^*$. Each agent's surplus $\pi_i = q_i - c_i(a_i^*)$. Let Θ_i denote the set of all the possible output level with unilateral deviation by agent i, namely $\Theta_i = \{\theta | \exists \hat{a}_i \text{ such that } \theta = \mathcal{F}(\hat{a}_i, a^*_{-i})\}$. Since we are focusing on Nash equilibrium where only unilateral deviation will be considered in incentive provision. For any realization of output level $\hat{\mathcal{F}} \notin \cap\Theta_i$, we can easily find a balanced sharing rule that punish some agents that might have caused the outcome as severe as possible. Thus we would focus our attention to the sharing rule for output level $\mathcal{F} \in \cap\Theta_i$.

Let $\Theta = \cap\Theta_i$, for each $\theta \in \Theta$, let $a_i(\theta, \mathbf{a}^*)$ denote the effort agent i may have exerted to produce the output level θ given other agents' effort as $\mathbf{a}^*_{-i}$. We have $\mathcal{F}(a_i(\theta, \mathbf{a}^*), \mathbf{a}^*_{-i}) = \theta$. Similarly we define the possible effort cost for agent i as $v_i(\theta, \mathbf{a}^*) = c_i(a_i(\theta, \mathbf{a}^*))$. The condition for implementability of $\mathbf{a}^*$ can be summarized as the following theorem:

Theorem 4.6. $\mathbf{a}^*$ *can be implemented if and only if*

$$\rho(\mathbf{a}^*) \leq \frac{1}{1-\delta}[\mathcal{F}^* - \sum_{i=1}^{n} c_i(a^*)] - \frac{\delta}{1-\delta}\sum_{i=1}^{n}\bar{\pi}_i \tag{4.8}$$

, *where* $\rho(\mathbf{a}^*) = \max\{\theta - \sum_{i=1}^{n} v_i(\theta, \mathbf{a}^*)\}$

Proof. Necessity: If $\mathbf{a}^*$ can be implemented, there exist a sharing rule $s(.) = (s_1(.), \ldots, s_n(.))$ such that

$$q_i - c_i(a_i) \geq s_i(\theta) - v_i(\theta) + \frac{\delta}{1-\delta}(\bar{\pi}_i - \pi_i) \tag{4.9}$$

holds for all θ and i. Hence for each θ, summing all the inequality we have

$$\theta - \sum_{i=1}^{n} v_i(\theta, \mathbf{a}^*) \leq \frac{1}{1-\delta}\sum_{i=1}^{n}\pi_i - \frac{\delta}{1-\delta}\sum_{i=1}^{n}\bar{\pi}_i \tag{4.10}$$

The inequality thus holds for ρ as well.

Sufficiency: If $\rho(\mathbf{a}^*) \leq \frac{1}{1-\delta}[\mathcal{F}(a_i', \mathbf{a}^*_{-i})) - \sum_{i=1}^{n} c_i(a_i')] - \frac{\delta}{1-\delta}\sum_{i=1}^{n}\bar{\pi}_i$, for each θ, we have inequality 4.10, rearranging the terms we have:

$$[\mathcal{F}^* - \sum_{i=1}^{n} c_i(a^*)] - [\theta - \sum_{i=1}^{n} v_i(\theta, \mathbf{a}^*)] \geq \frac{\delta}{1-\delta} \sum_{i=1}^{n} (\bar{\pi}_i - \pi_i) \tag{4.11}$$

Thus for each θ, there exist balanced sharing $s(\theta) = (s_1(\theta), \ldots, s_n(\theta))$, such that for each i

$$q_i - c_i(a^*) \geq s_i(\theta) - v_i(\theta, \mathbf{a}^*) + \frac{\delta}{1-\delta}(\bar{\pi}_i - \pi_i) \tag{4.12}$$

Thus, IC constraints hold with sharing rule $s(.)$ and $\mathbf{a}^*$ can be implemented □

Theorem 4.6 provides a necessary and sufficient condition for an effort level to be implemented by a relational contract with abstract sharing rules. However this only holds under general case where no restriction is put on the functional form of the sharing rule. The flexibility enables us to find proper values for $s_i(\theta)$ across all i and θ. The result thus holds for any distribution of final output $(q_1, \ldots, q_n)$, where for each i, $q_i - c_i(a_i^*) \geq \bar{\pi}_i$.

4.5 Linear sharing rule

As a general assumption throughout Rayo [2007]'s model, linear sharing is the most commonly seen sharing in real life because of its simplicity. However, whether linear sharing can be optimal is not clear. The restriction makes our search for optimality much easier but the comfort might be on the cost of efficiency. Thus in this section we would try to disclose whether anything is lost by assuming linearity in the sharing rule.

From now on we would assume linear sharing and search for the optimal linear sharing to see if there is any efficiency loss by putting linearity restriction on the sharing rule. For each i, we assume $s_i(\mathcal{F}) = \alpha_i \mathcal{F} + \beta_i$ and $\sum \alpha_i = 1$, $\sum \beta_i = 0$.

In general sharing, we need to specify shares for each agent and each output level. However for each individual we only need to solve two parameters. With linear sharing constraint 4.7 would be affected by both parameter α and β, the utility among agents are no-longer transferable. Any constant transfers of output will affect individual IC constraints. Thus, the maximization of total team surplus can only support certain distribution outcomes under linear sharing. In this sense if we restrict to linear sharing rules, maximizing team surplus is merely a typical case with a specific social welfare function. While there might exist other outcomes that some specific agents might prefer while don't maximize the team surplus. Which social welfare function the team would choose depends on the organization structure, it might be principal-agent structure where there is one agent who dominates and put his own welfare as the priority, or it could be a Nash bargaining process. We raise this issue here to make critical argument about the surplus maximization in Rayo [2007]. When individuals' surpluses enter their IC constraints endogenously, utilities are no longer transferable. Which specific target effort will be preferable depends on the social welfare function of the team, thus we will leave the efficiency issue to the future study and focus on solving the optimal contract first.

Due to the budget balancing constraint, it's impossible to increase every agent's share to 1. The problem is then how to derive an optimal sharing rule to create the maximum surplus possible. We start by simplifying our IC constraints

Lemma 4.7. *Given* $\mathbf{a}$*, IC constrains can hold for* i *if and only if*

$$\frac{\mathcal{F}(\mathbf{a}) - \sum_{i=1}^{n} c(a_i)}{1-\delta} \geq \sum_{i=1}^{n} \sup[s_i(\mathcal{F}(a_i', \mathbf{a}_{-i})) - c_i(a_i')] + \sum_{i=1}^{n} \frac{\delta}{1-\delta}\bar{\pi}_i \tag{4.13}$$

Proof. See Appendix B □

Lemma (4.7) shows that instead of looking into the IC constraints agent by agent, we can focus our attention on one constraint (4.13) and there always exist balanced transfers on the constant payment β to make all individual IC constraints (4.6) hold. In constraint (4.13), with the target effort level $\mathbf{a}$ given, the only term left that would be varying with the sharing rule is $\sum_{i=1}^{n} \sup[s_i(\mathcal{F}(a_i', \mathbf{a}_{-i})) - c_i(a_i')]$, which is the key to find the optimal sharing rule.

However, the procedure of searching for the optimal linear sharing rule needs to be based on a given target effort level. It's of interest to see under which conditions, first-best outcome can always be implemented with any linear sharing rule. We have the following theorem:

For any agent i, giving the sharing rule s and target effort level $\mathbf{a}^*$, let a_i' denote agent i's best static response without relational incentive, then a_i' must solve

$$\max_{a_i} s_i(a_i, a_{-i}^*) - c_i(a_i) \tag{4.14}$$

For each agent i, given his peer effort level 'a^*_{-i} the relation between the agent's best respond effort level a'_i and his share of output α_i is described by the Nash equilibrium conditions. For given target effort level $\mathbf{a}$, the optimal sharing rule should try to satisfy IC constraint (4.13)

Thus the optimization problem becomes:

$$\min_{\alpha_1,\ldots,\alpha_n} \sum_{i=1}^{n} \{[\alpha_i \mathcal{F}(a'_i(\alpha_i), a_{-i}) - c(a'_i(\alpha_i))] \tag{4.15}$$

Subject to: $\sum_{i=1}^{n} \alpha_i = 1$.

Since the RHS of each constraint (4.7) involves a maximization problem, where $a'_i \in \arg\max[s_i(\mathcal{F}(a'_i, \mathbf{a}_{-i})) - c_i(a'_i)]$, we need to show that our objective function is differentiable with respect to α_i for all i

Lemma 4.8. *Given* $\mathbf{a}$, *for each agent* i *, the mapping* $\alpha_i \to a'_i$ *is a function differentiable on* $\mathbb{R}^+$

Proof. See Appendix B □

With lemma (4.8), we are able to analyze the impact of the sharing rule on IC constraint through calculus.

Lemma 4.9. *Given* $\mathbf{a}$, *for each* i *the incentive power is increasing in his own share* α_i.

Proof. See Appendix B □

Lemma 4.9 shows how the share α can influence the IC constraint and result is very straightforward: given the amount of implicit incentive available, the higher the explicit incentive, the more likely the effort level is implementable. However, due to the budget balancing condition, we cannot increase every agent's share to 100%. The optimal linear sharing would thus provide as much incentive power as possible to fill in the "incentive gap" left by the relational incentive. And we would show in the following theorem that there is a best way to allocate agents' shares.

Theorem 4.10. *The optimal linear sharing rule s^* given target effort level $\mathbf{a}^*$ satisfies the following property: For each agent i, with the optimal share α_i^*, the optimal deviation output level must be the same, such that $\mathcal{F}(a_1', a_{-1}^*) = \cdots = \mathcal{F}(a_i', a_{-i}^*) = \cdots = \mathcal{F}(a_n', a_{-n}^*) = \bar{\mathcal{F}}$*

Proof. See Appendix B □

Given the property that the implicit incentive can be shifted around agents without fraction, the optimal sharing rule tends to minimize the sum of deviations among all agents. Theorem 4.10 would make each agent deviate to the same output level given his colleagues' contribution fixed. This comes from the fact that if an agent maximize it's own payoff given others effort level, then the marginal effect of changing this single agent's share would be equal to the deviation output level at the current share. Thus minimum is achieved when this marginal effect is the same for everyone, providing no chance of further reducing the sum of deviation by shifting shares among agents.

Up to this point we have shown the property of the optimal linear sharing rule, but there might exist other non-linear or abstract sharing rules that can achieve a better outcome. Thus we need to extend our analysis to general sharing rules.

We would start by relaxing the linear constraint to consider other differentiable sharing. Let $\hat{\theta}$ denote the deviation output level, we would show that $\hat{\theta}$ is uniquely defined for any differentiable sharing rule.

Lemma 4.11. *Given any differentiable sharing rule* $S = \{s_1(.), s_2(.), \ldots, s_n(.)\}$, *if* $\theta' \arg\max\{s_i(\theta) - v_i(\theta))\}$ *for all* i, $\theta' = \hat{\theta}$.

Proof. See Appendix B □

The lemma shows that under differentiable sharing, if we want each agent's optimal deviation output θ to be the same, then the deviation output is unique. Unlike linear sharing, where the marginal change in share is constant for any output level and has to sum up to 1 across all agents, non-linear sharing can have the flexibility of having the marginal effect been summed up to other constants. Although with the lemma non-linear sharing would be equivalent to linear sharing if the optimal deviation is specified at $\hat{\theta}$. We need to further compare with other cases when each agent's optimal deviation output levels are different.

Theorem 4.12. *For any balanced differentiable sharing rule* $\{s_1(.), s_2(.), \ldots, s_n(.)\}$, *we have* $\sum_{i=1}^{n}[s_i(\theta_i) - v_i(\theta_i)] \geq \hat{\theta} - \sum_{i=1}^{n} v_i(\hat{\theta})$, *where for each* i, $\theta_i \arg\max\{s_i(\theta) - v_i(\theta_i)\}$.

Proof. Since for each i, $\theta_i \arg\max\{s_i(\theta) - v_i(\theta_i)\}$. We have $s_i(\theta) - v_i(\theta_i) \geq s_i(\hat{\theta}) - v_i(\hat{\theta})$.

Summing all the inequality across all agent i, we have $\sum_{i=1}^{n}[s_i(\theta_i)-v_i(\theta_i)] \geq \sum_{i=1}^{n}\{s_i(\hat{\theta})-v_i(\hat{\theta})\}$. Since the sharing rule is balanced, $\sum_{i=1}^{n}\{s_i(\hat{\theta}) = \hat{\theta}$. The proof is complete. □

Given our objective function as minimizing the sum of deviation benefits, theorem 4.12 shows that there does not exist any other differentiable sharing rule that can give a better outcome than the optimal linear sharing. The inequality in 4.12 holds with equality if and only if for all i, $\theta_i = \hat{\theta}$.

Example 4.2. *Given a target effort level let $g_i(.) : a_i \to \mathcal{F}(a_i, a_{-i})$. If for all $i \in N$, $g_i(.) = g(.)$, the optimal sharing rule must be symmetric for all agent, with all agent getting the same proportion of share.*

Given their peer effort level, if each agents' roles in the production are homogenous, by theorem 4.10, the optimal sharing rule must specify the same deviation output level for all agents, thus the share for each agent must be the same.

Example 4.3. *For a production technology, if for all i and any arbitrary a_{-i}, $\mathcal{F}(0, a_{-i}) = 0$, then all agents must receive strictly positive shares.*

In the special case where every agent is indispensable in team production, they must receive at least some incentive from the sharing to exert positive effort.

The above two examples show how the optimal linear sharing rule is related to the production structure and the target effort level. In theorem 4.10 we investigated the optimal sharing rule treating target effort level given. But whether the self-enforcing constraint is fulfilled remains to be solved. The relation between the optimal linear sharing and optimal general sharing are not clear by far. We would focus on comparing these two in the next section.

4.6 General verses linear sharing

In section 4.4, we showed that an effort vector $\mathbf{a}^*$ and sharing of output $(p_1, \ldots, p_n)$ can be implemented if and only if $\rho(\mathbf{a}^*) \leq \frac{1}{1-\delta}[\mathcal{F}^* - \sum_{i=1}^{n} c_i(a^*)] - \frac{\delta}{1-\delta}\sum_{i=1}^{n} \bar{\pi}_i$. Once the condition hold any outcome $(p_1, \ldots, p_n)$ that satisfies $p_i \geq \bar{\pi}_i$ can be implemented.

While for linear sharing, from the previous section the optimality specifies every agent's optimal deviation as $\hat{\theta}$, thus the necessary and sufficient condition has to be specific for each agent i such that:

$$\alpha_i\hat{\theta} + \beta_i - v_i(\hat{\theta}) \leq \frac{1}{1-\delta}[\alpha_i\mathcal{F}^* + \beta_i - c_i(a_i^*)] - \frac{\delta}{1-\delta}\bar{\pi}_i \tag{4.16}$$

It's clear that with linear sharing, the validity of each of these individual constraints depends on both α_i and β_i. Unlike the general sharing, the distribution of the outcome would not affect constraint 4.8. While for linear sharing, each individual constraint will be affected through the alteration of $beta_i$s, thus we do not have the freedom to choose the distribution of output.

If we put the problem output distribution aside, the question we are interested in is: if effort $\mathbf{a}^*$ can be implemented by general sharing, can it also be implemented by linear sharing with restrictions to the distribution of the output? The answer lies within the inner relation between the two types of sharing.

Let $\theta^* \in \arg\max\{\theta - \sum_{i=1}^{n} v_i(\theta)\}$, the key question of comparison of the general sharing with linear sharing lies in the comparison of the θ^* specified by general sharing and the $\hat{\theta}$ specified by linear sharing.

Lemma 4.13. $\theta^* = \hat{\theta}$

Proof. See Appendix B □

From lemma 4.13, we can see that the optimal deviations specified by general sharing and linear sharing are the same. Thus the condition 4.8 with general sharing can be treated as a summation of the n conditions specified by linear sharing. With linear sharing, effort level $\mathbf{a}^*$ can be implemented if and only if the n conditions holds simultaneously and the choose of $beta_i$s is important. We will show in the following theorem the existence of *beta* that satisfies the individual constraints with α.

Theorem 4.14. *If* $\mathbf{a}^*$ *can be implemented by general sharing, then there exist some linear sharing rules that can implement* $\mathbf{a}^*$ *with some* β*s.*

Proof. See Appendix B □

Note that the above theorem talked about the implementation of effort only. Due to the simple nature of linear sharing, i.e. only two variables need to be fixed; we are able to implement some effort in the same way as general sharing but with limitations on the distribution of the outcome. For example, if condition 4.8 binds with strict equality, then the vector β is unique and only one of the outcomes that can be implemented by general sharing can be implemented by linear sharing.

In this sense general sharing always dominates linear sharing. We can assume linear sharing without loss of generality if and only if the social welfare function within the team is the team surplus. We will use a numerical example to show the difference of the two.

We would take a numerical example to further illustrate our result.

Example 4.4. *Consider a team consists of agents 1, 2 and 3. Let* $\mathbf{a} = \{a_1, a_2, a_3\}$*, with the output* $\mathcal{F} = 2a_1^{\frac{1}{2}}a_3^{\frac{1}{2}} + 2a_2^{\frac{1}{2}}a_3^{\frac{1}{2}}$*. Each agent's cost function is defined as follow:* $c_1 = a_1^2/2$*,* $c_2 = a_2^2/2$*,* $c_3 = a_3^2$*. Suppose the sub-team structure satisfies the condition in theorem 4.4. Let* $\delta = \frac{1}{2}$*.*[3]

In the above example, the first best effort level is $\mathbf{a}^* = \{1, 1, 1\}$, generating output $\mathcal{F}^* = 4$ and surplus $\mathcal{F}^* - \sum_{i=1}^{n} c_i(a_i^*) = 2$. Figure 4.1 plots $\theta - \sum_{i=1}^{n} v_i(\theta)$ against $\theta \in [2, 4]$. From the graph, the maximum point 2.66 which is the ρ given $\mathbf{a}^* = \{1, 1, 1\}$ as the target effort level. According to theorem 4.6, the first-best effort can be implemented by relational contract if and only if $\sum_{i=1}^{n} \bar{\pi}_i \leq 1.34$.

Also theorem 4.14 suggests that if first-best outcome can be implemented, it can be implemented with linear sharing. Figure 4.2 plots the amount of implicit incentive needed for all possible non-negative sharing. The minimum is obtained when linear sharing specifies the shares as $\alpha = \{0.235, 0.235, 0.53\}$. Since agent 1 and 2's roles are perfectly symmetric in this setting, the optimal linear sharing would give them the same share. Theorem 4.14 proves that the minimum requirement for implicit incentive under linear sharing is exactly the same as ρ. However, when using linear sharing, some distribution of output levels cannot be supported. Given α and from

[3] A static case of this example can be found in Nandeibam [2002].

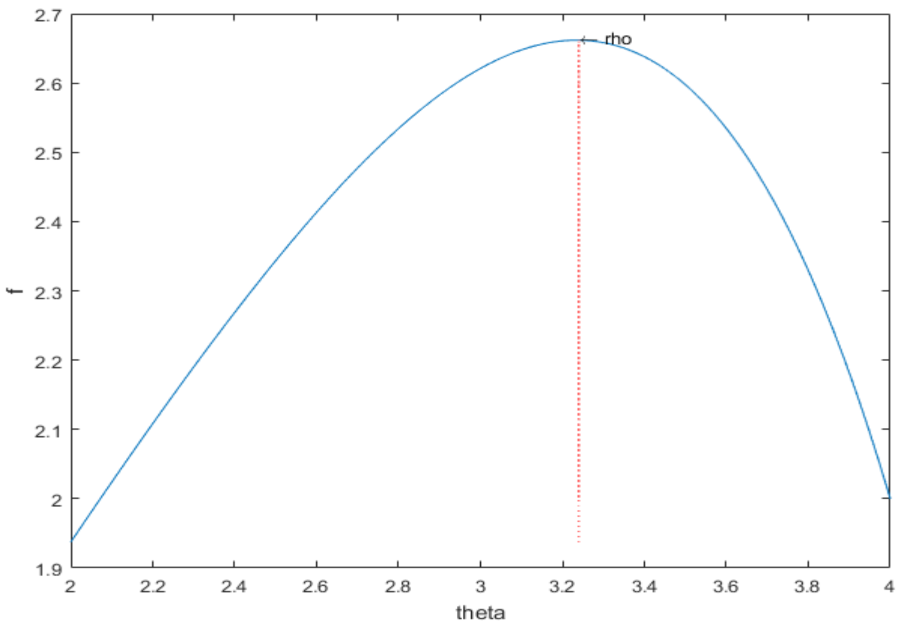

FIGURE 4.1: Searching for ρ

the individual IC constraints we have for agent 1 and 2: $\frac{1}{1-\delta}[0.235*4+\beta-0.5] \geq [0.235*(2\sqrt{0.765}+2)+\beta-\frac{0.765^2}{2}]+\pi_i$, where 0.765 is agent 1 and 2's optimal effort level with sharing rule alone. we have $\beta_1 \geq \bar{\pi}_1 - 0.3$ and $\beta_2 \geq \bar{\pi}_2 - 0.3$. The same reasoning with agent 3 we have $\beta_3 \geq \bar{\pi}_3 - 1$. Thus if we try we implement the first-best outcome with linear sharing rule, each agent's surplus must satisfy $\pi_1 \geq \bar{\pi}_1 + 0.15$, $\pi_2 \geq \bar{\pi}_2 + 0.15$ and $\pi_3 \geq \bar{\pi}_3 + 0.11$. While for general sharing, we only need $\pi_i \geq \bar{\pi}_i$ for each i.

We can easily visualize the difference in figure 4.3, where an arbitrary point O inside the big triangular represents a possible surplus distribution. Here we normalize the outside options to 0. The distance O to each side of the big triangular represents the surplus distributed to each agent. The shadowed area shows the implementable outcome with linear sharing, while the general sharing can implement both white and shadowed area. It's clear that with linear sharing, it's impossible to concentrate all the surplus to one agent and leaving the rest of the team getting their outside options.

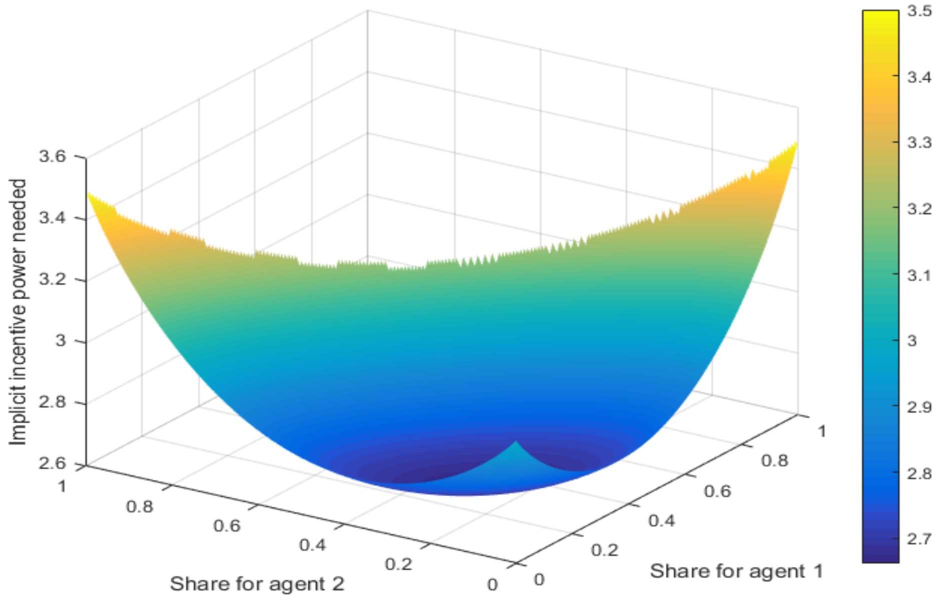

FIGURE 4.2: Minimum implicit incentive needed with different linear sharing

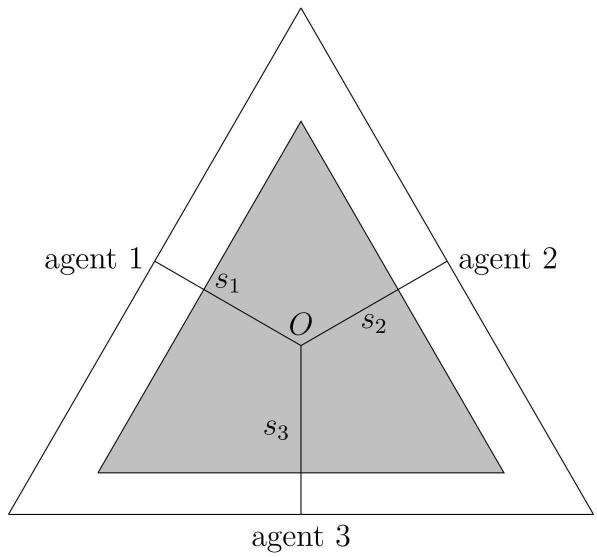

FIGURE 4.3: Distribution of surplus under general and linear sharing

4.7 Conclusion

This chapter has studied team relational contract under with deterministic sub-team performances. The two main questions we tried to answer are: firstly, how can the information on sub-team performance be utilized to the maximum in relational contracting (through the grouping mechanism that allocate agents to different sub-teams); Secondly where the efficiency lies and what's the role that sharing rules can play.

In answering the first question we decompose the searching of the optimal relational incentive into two steps. In the first step, we have shown that there exists a simple grouping mechanism such that any single deviation from any agent will be perfectly detected within a proper subset of the agent set. This ensures that anyone who deviates will be punished to the maximum while there always exist some other agents to balance the budget. In the second step, we measure the maximum amount of punishment can be enforced through the self-enforcing constraint.

The optimal implicit incentive suggests that the incentive power from relational incentive only depends on the surplus of the team on the equilibrium path. Given a target effort, we show the necessary and sufficient condition for a target effort vector can be supported with the combination of balanced sharing rule with relational contract. Under general setting, agent's utilities are transferable, namely any distributional objective can be implemented.

However, under linear sharing rule, we are only able to implement the efficient effort vector under certain range of distribution of surplus. This is due to the fact that only two parameters are involved in determining an agent's share. Once the optimality fixed the slope of the linear function, we can no longer change the constant freely, resulting in the failure of implementation of some outcomes.

Chapter 5

A Bayesian Approach to Linear Public Good Games

5.1 Introduction

In previous chapters we study the team incentives in monetary terms, however experimental studies have shown that humans do not always behave in a rational way to maximize their monetary payoffs. In the dictator game investigated by Forsythe et al. [1994], the average gift sent out by dictators is 20%. Fehr et al. [1993] also shows that people tries to be behave nicely to others at their own costs and expect the same in return. All these experiments suggest that good outcome can be achieved without monetary incentives and people's social preferences play the key role in people's counter-theoretical behaviors. In this chapter we would give some preliminary analysis on the theory of inequality aversion, in order to build up the base for future theoretical and experimental analysis.

Although we do find people's preferences towards equal outcome, people will not always have enough information to convince themselves to pursue the "good" equality. Especially when their decision have to be made upon strict assumptions upon others. A vivid example is established by Fehr and Schmidt [1999], through comparing the outcome of linear public good experiments with and without punishment. When subjects are asked to contribute towards a linear public good simultaneously, more than a half chooses to contribute nothing. However, when punishment is allowed after the result is revealed, the majorities contribute fully towards the public good. The results suggest that people do care about equality: If they are selfish, no one should choose to punish free-riders at his own cost, when this punishment cannot change the outcome of the public good game. Moreover, it suggests that people also believe other people are averse to inequality: convinced that free-riders will be punished in the second stage, most subjects contribute fully.

Two key questions arises in the theory explanation of Fehr and Schmidt [1999]'s model. First, Fehr and Schmidt [1999] assumes the subjects' preferences over inequality are public information. This is a very strong assumption given the fact that most subjects shouldn't know each other. Their result heavily relies on this assumption. In this chapter we would assume incompleteness of information, by the preference parameter is known only to each subject themselves, but the distribution of the parameters are publicly known. We would derive the Bayesian

Nash equilibriums of the linear public good game without punishment and have a basic discussion about the punishment scheme adopted by Fehr and Schmidt [1999].

The second question is: under the game without punishment, players would not know the exact contribution vector, however, Fehr and Schmidt [1999]'s inequality aversion is still calculated based on the contribution vector. In fact, when players are only able to observe the final output of the public good, the only inequality they are able to derive is from the difference between their own payoff with the average. We derived the theory to consider this situation and the equilibrium behavior suggest that no matter a player is averse to inequality or not, if he believes there are people who doesn't care about inequality, he would not cooperate. This explains the high proportion of 0 contributions in linear public good games without punishment, but further experiments need to be conducted to test this idea.

In general this chapter aims to build the theory foundation for inequality aversion that can be applied in incentive provision in team production. We try to develop the theory base on Fehr and Schmidt [1999]'s framework and shed light on some future experimental studies.

5.2 Linear public good model and inequality aversion

We consider a linear public good model where individuals care about inequality. Suppose there are n agents taking part in a project where the payoff of the project is linear towards the sum of all agents' contributions. We denote $g_i \in \mathbb{G}$ as agent i's contribution and the action space $\mathbb{G}$ is a continuum identical to all agents. Let the final payoff function be $\mathbf{F} = a\sum_{i=1}^{n} g_i$, with $a < n$. Let $\mathbf{S}$ be the sharing rule among these agents and s_i describes the specific share for agent i. For each i, $0 < s_i < 1$ and $\sum_{i=1}^{n} s_i = 1$.

Through out the chapter we assume each agent not only care about his own payoff but also the inequality between each other. The modeling of inequality aversion has two key elements: the measurement of inequalities based on publicly observable information and the mapping relationship between the degree of inequality and the agents' averse feeling towards the inequality. The measurement of inequality τ can vary from different public good games and will be discussed specifically in each game setting. Let $\tau_{ij} = x_i - x_j$ be a binary inequality measure on x[1] for agent i with agent j, we would take Fehr and Schmidt [1999]'s approach by assuming players' preferences toward inequality are linear. For each i we have,

$$f_i(\tau_{ij}) = \begin{cases} \beta_i \tau_{ij} & \text{if} \tau_{ij} \geq 0, \\ \alpha_i \tau_{ij} & \text{if} \tau_{ij} < 0 \end{cases} \tag{5.1}$$

[1]Without loss of generality we assume agents' utility functions are increasing on x.

The inequality aversion structure in 5.1 states that subjects cares both advantageous inequality and disadvantageous inequality. The degree of aversion/preference towards advantageous inequality and disadvantageous inequality are controlled by parameter β_i and α_i respectively. Although the linear assumption might not be fully realistic, it provides us an simplest possible structure to look into public good games while being able to explain a lot of results in the public good game experiments.

In standard linear public good game, the only inequality the subjects will face would be the difference between the net profit through contributing to the public good. Let $x_i = \frac{a}{n}\sum_{i=1}^{n} g_i - g_i$, $\tau_{ij} = x_i - x_j = g_j - g_i$. If $n = 2$, subject i's utility function can be expressed as :

$$U_i(g_i, g_j) = \frac{a}{n}\sum_{i=1}^{n} g_i - g_i - \beta_i \max\{g_j - g_i, 0\} - \alpha_i \max\{g_i - g_j, 0\} \tag{5.2}$$

If $n > 2$, then each player needs to compare himself with other $n-1$ players, we thus normalize the aversion towards inequality by taking the average across all the $n-1$ comparisons. This ensures the relative impact of inequality aversion is invariant to the number of players in the game. When $n > 2$, we have:

$$U_i(g_i, g_{-i}) = \frac{a}{n}\sum_{i=1}^{n} g_i - g_i - \frac{\beta_i}{n-1}\sum_{i\neq j} \max\{g_j - g_i, 0\} - \frac{\alpha_i}{n-1}\sum_{i\neq j} \max\{g_i - g_j, 0\} \tag{5.3}$$

5.3 Incomplete information

Fehr and Schmidt [1999]'s theory assumes that agents' utility functions are heterogeneous in the sense that parameters α and β may vary across agents, and the equilibrium behaviors are derived under the assumption that α_is and β_is are publicly known to everyone. However it is a very strong and somehow unrealistic assumption, especially under circumstances where subjects taking part in the experiments are randomly chosen.

In this chapter we would propose a different approach towards the game and try to explain the results of some previous findings from experiments. We relax the assumption of complete information by assuming the parameters α_is and β_is are privately known. While β_is are independent and identically distributed over $(-\gamma, 1)$ with density function $f(\beta)$. Here we assume independent distributions because we believe subjects' beliefs over β are formed over a much bigger sample within some certain cultural environment rather than the small sample participating the experiment. Thus we do not provide the prior joint distribution over $(\beta_1, \ldots, \beta_n)$, as knowing his/her own β_i would not help agent i to infer other's β.

Different β shows different preferences towards advantageous inequality:

- If $-\gamma < \beta < 0$, then subject would gain strictly positive utility from being better off than others. The lower bound γ represents the maximum cost such agent would be willing to pay to increase his advantageous inequality with others.

- If $0 \leq \beta < \frac{1}{2}$, then subject would have disutility from being better off than others, while this feeling is not strong enough for him/her to transfer some money to others to get equal outcome.

- If $\frac{1}{2} \leq \beta < 1$, then subject's aversion towards advantageous inequality would enable the subject transforming his money to others to pursue an equal outcome. Here we rule out the case that $\beta \geq 1$, as we assume that no subject would throw money away to get a perfect equal outcome.

Under linear public good setting, the mechanism of monetary transfer is based on the contribution to the public good, given $g_i < g_{-i}$, a marginal increase in g_i would decrease i's monetary payoff by $\frac{n-a}{n}$, while the average inequality between agent i with other agents would decrease by 1. Thus in a two player linear public good game, if subject i's contribution is less than subject j, he would only increase his contribution if and only if $\beta_i \geq \frac{a-n}{n}$.

For parameter α which controls subject's preference towards disadvantageous inequality, we would assume that for each i, $\alpha_i \geq \max\{0, \beta_i\}$. The assumption captures the following ideas: firstly, agents are always averse to disadvantageous inequality; secondly, if an agent is averse to both types of inequality, he would prefer advantageous inequality to disadvantageous inequality. Given β_i, α_i is distributed over $[\max\{0, \beta_i\}, \lambda]$, with conditional density function $h(\alpha_i|\beta_i)$. To interpret the meaning of the upper bound λ, let's consider a case where player 1 offers player 2 an opportunity to join a project where player 2 need to decide how much he wish to invest. Suppose as the project leader player 1 always gets more than player 2, and the net marginal payoff for player 2 from investing is always positive and we normalize it to 1, while the marginal payoff for player 1 from player 2's investment is x and $x > 1$. If player 2 doesn't care about inequality he would always participate in the project, as it's beneficial to him. However, he would reject the offer if $\alpha_2 > \frac{1}{x-1}$. By setting an upper bound for α we would rule out the extreme case where agents would sacrifice their own benefits to avoid just tiny amounts of disadvantageous inequality.

5.4 Bayesian Nash equilibrium in public good game

Consider $n > 2$ players participating a public good game, each agent i has an initial endowment y that can be invested into the public good. At the beginning of the game, each agent simultaneously choose a contribution level $g_i \in [0, y]$. The monetary payoff of the public good $\mathbf{F}$ is linearly dependent on the sum of contributions gathered from the players, with $\mathbf{F} = a \sum_{i=1}^{n} g_i$, where a is a parameter that can be controlled for during experiment and $a < n$. After the contributions are gathered, the public good will be equally shared among all players, and the anonymous contribution vector would be revealed.

Each player's preference towards inequality is characterized by a two-dimensional type vector $\theta_i = [\alpha_i; \beta_i]$. The utility function for each player i is:

$$U_i(g_i, g_{-i}, \theta_i) = y + \frac{a}{n}\sum_{i=1}^{n} g_i - g_i - \frac{\beta_i}{n-1}\sum_{i \neq j} \max\{g_j - g_i, 0\} - \frac{\alpha_i}{n-1}\sum_{i \neq j} \max\{g_i - g_j, 0\} \tag{5.4}$$

The equilibrium behavior is described by the following theorems:

Theorem 5.1. *If $\beta_i < 1 - \frac{a}{n}$, with any $\alpha_i \geq |\beta_i|$, agent i with θ_i has a dominant strategy to contribute $g_i = 0$*

Proof. See Appendix C □

Theorem 5.1 shows that if the degree of aversion towards advantageous inequality cannot off set the loss in monetary payoff through contribution, players would contribute 0 no matter what others contributions are. Even if other players contribute some positive amounts, inequality aversion would not help to provide incentive to players who doesn't care about inequality that much. While for players with β higher than $1 - \frac{a}{n}$, whether there exist any positive contributions needs further clarification.

Suppose Θ is the set such that if $\theta_i \in \Theta$, positive contribution is the best respond to each other. Given all other players j such that $\theta_j \in \Theta$ contribute strictly positive amount $g > 0$. The payoff for player i contributing $0 \leq g_i \leq g$ given the number of contributors k is:

$$U_i(k) = y + \frac{ak}{n}g + (\frac{a}{n} - 1)g_i - \frac{\alpha_i(n-1-k)}{n-1}g_i - \frac{\beta_i k}{n-1}(g - g_i) \tag{5.5}$$

, where $0 \leq k \leq n-1$

Let $\mathbf{p} = p(\theta \in \Theta)$, the probability we assign to each realization of k is $C_{n-1}^k \mathbf{p}^k (1 - \mathbf{p})^{n-1-k}$, the expected payoff function for player i can be written as:

$$
\begin{aligned}
E[U_i(g_i)] =& \sum_{k=0}^{n-1} C_{n-1}^k \mathbf{p}^k (1-\mathbf{p})^{n-1-k} U_i(k) \\
=& y + \frac{a}{n}(n-1)\mathbf{p}g + (\frac{a}{n} - 1)g_i - \alpha_i(1-\mathbf{p})g_i - \beta_i \mathbf{p}(g - g_i)
\end{aligned}
$$

It's easy to show that player i would never contribute higher than g. The marginal impact of increasing g_i can be interpreted in the following way: first of all, a marginal increase oin g_i by ε would decrease the aversion towards expected advantageous inequality by $\beta_i \mathbf{p}\varepsilon$, on the other hand it incurs a monetary cost $(1 - \frac{a}{n})\varepsilon$ and an increase in the aversion towards the expected disadvantageous inequality $\alpha_i(1-\mathbf{p})\varepsilon$. Since $\theta_i \in \Theta$, $\beta_i \mathbf{p} \geq 1 - \frac{a}{n} + \alpha_i(1-\mathbf{p})$ must hold. Rearranging in inequality we have:

$$
\mathbf{p} \geq \frac{\alpha_i + 1 - a/n}{\alpha_i + \beta_i} \tag{5.6}
$$

Since 5.6 must hold for any $[\alpha_i; \beta_i] \in \Theta$, $\mathbf{p} \geq \sup\{\frac{\alpha_i + 1 - a/n}{\alpha_i + \beta_i}\}$. Since $\beta_i \geq 1 - \frac{a}{n}$, the supremum is obtained under $\inf \beta$ and $\sup \alpha$, given the property that $\alpha_i \geq \beta_i$ for all i, the problem is to find x and ys such that: $\Theta = \{[\alpha; \beta] || x \leq \beta \leq alpha \leq y\}$ and

$$\mathbf{p} \geq \frac{y+1-a/n}{y+x} \tag{5.7}$$

, where $\mathbf{p} = \int_x^y \int_\beta^y h(\alpha|\beta) f(\beta) d\alpha d\beta$

Remark 5.2. If any Bayesian Nash equilibrium with positive contribution exist, x must be strictly greater than $1 - \frac{a}{n}$

It's an obvious implication of the equilibrium structure of our game, that when players are evaluating the expected number of players who are going to contribute positive amount, players with $\beta = 1 - \frac{a}{n}$ will only contribute unless he think everyone is going to contribute and of course this is not possible. In general only players with high enough β and low enough α will be able to be incentivized through inequality aversion.

The equilibrium behavior can be summarized as the following theorem:

Theorem 5.3. *If* $\max\{\int_x^y \int_\beta^y g(\alpha|\beta) f(\beta) d\alpha d\beta - \frac{y+1-a/n}{y+x}\} < 0$*, for any type of agent* θ_i*, there is a unique equilibrium of contributing 0.*

Otherwise, there exist Bayesian Nash equilibriums such that for each player i*, if* $\theta_i \in \Theta = \{[\alpha;\beta] | x \leq \beta \leq \alpha \leq y\}$*, where* $\int_x^y \int_\beta^y g(\alpha|\beta) f(\beta) d\alpha d\beta = \frac{y+1-a/n}{y+x}$*, then* $g_i = g \in [0, y]$*; if* $\theta_i \notin \Theta$*, then* $g_i = 0$

Proof. See Appendix C □

When players are informed about each others preferences over inequality, Fehr and Schmidt [1999] divided the players into two groups, one group as the free riders who have a dominant strategy to contribute 0 and all the other agents are grouped into the cooperators. And the existence of equilibrium such that cooperators would contribute depends on whether the proportion of free-riders is small enough and whether the cooperators' αs are low enough. The equilibrium of our model shares a very similar structure; we admit the existence of strict free-riders who will free-ride on others contribution no matter what happens, while there might exist cooperators who are willing to cooperate within themselves. The existence of cooperators depends on the following conditions: first, any of such cooperators must be averse enough towards free-riding while not been too upset about the free-riders; second, the probability of a player satisfying this property needs to be high enough. However, what Fehr and Schmidt [1999] has described is merely a typical case of our theorem, where the boundary of cooperators' α and βs are fixed. Theorem 5.3 shows that the definition for cooperators can vary, there might exist a set of combinations of x and y that satisfies the condition in the theorem. It would be interesting to investigate this issue in the future with some specific density functions of α and β.

5.5 Inequality aversion based on average

By far we have been strictly following Fehr and Schmidt [1999]'s definition of inequality. Fehr and Schmidt [1999]'s calculation of inequality is based on the contribution vector even if it is not revealed in the linear public good experiment without punishment. On one hand, individual's strategy is formed under some predictions of other's possible action; on the other hand, when the contribution vector is not revealed at the end of the game, it's not clear when making the decision, players would calculate his/her expected payoff based on individual contributions or simply the average contribution he/she observes. Hence in this section, we would discuss the equilibrium behavior under the other case when inequality aversion is based on public information only. Given the final output of the public good **F**, let $\bar{g}_i$ be the average contribution of all other players excluding player i

Each player's payoff function thus is defined as follows:

$$U_i(g_i, \bar{g}) = y + a\frac{n-1}{n}\bar{g}_i + \frac{a}{n}g_i - g_i - \beta_i \max\{\bar{g}_i - g_i, 0\} - \alpha_i \max\{g_i - \bar{g}_i, 0\} \quad (5.8)$$

It's not difficult to show that under such circumstances, for agents with $\beta_i < 1 - \frac{a}{n}$, not contribute is still the dominant strategy. If any Bayesian Nash equilibrium with positive contribution can be supported, the major difference in the equilibrium behavior lies in the following theorem:

Theorem 5.4. *Any positive contribution can be supported in a Bayesian Nash equilibrium if and only if* $p(\beta_i < 1 - \frac{a}{n}) = 0$

Proof. See Appendix C □

Theorem 5.4 shows that the equilibrium behavior would become completely different when changing the definition of inequality aversion. If the aversion towards inequality is calculated based on average, players i will only contribute positive amounts if and only if $\beta_i \geq 1 - \frac{a}{n}$ for all i. This is a much stricter assumption as we not only require cooperators themselves satisfy $\beta \geq 1 - \frac{a}{n}$, they also need to be convinced that everyone else cares about inequality at least to some extent. In general, if players are aware that there are players who don't care about inequality in the game, then the comparison of average will make the contribution converge to 0. In a way this accords to the experimental findings that on average 73% of subjects in standard linear public good game contribute 0.

When player's contribution vector is not revealed, the condition for the existence of positive contribution has to be very strong. Since people need to have good faith on everyone who participate. While if the contribution vector is known afterwards, players only need to judge whether the disutility from disadvantageous inequality will dominate other feelings. Thus it's not clear whether the increasing average contribution when introducing a punishment scheme by Fehr et al. [1997] is due to the punishment or revelation of the contribution vector.

5.6 Punishment scheme

In Fehr and Schmidt [1999]'s extension towards linear public good experiment, they introduced a punishment scheme such that the new game consists two stages: in the first stage, the game is identical to the previous game; in the second stage, the contribution vector $(g_1, \dots, g_n)$ is revealed and each player i can simultaneously impose a punishment on other players $p_i = (p_{i1}, \dots, p_{in})$ where $p_{ij} \geq 0$ denote the punishment player i impose on player j. Each unit of punishment imposed by player i incurs a cost to player i himself, which we denote by c and $0 < c < 1$.

Since Fehr and Schmidt [1999] assumes $\beta_i \geq 0$ for all player i, no player would punish others just to make themselves better off than others.

Fehr et al. [1997]'s experimental result shows that about 80% of the subjects contributed fully in the first stage, and free-riders did face a vast punishment from majority of contributors in the second stage. Fehr and Schmidt [1999] provides the theory explanation of how full cooperation can be sustained as a sub-game perfect Nash equilibrium. This requires: first of all, a norm specifying how much players should contribute if they behave cooperatively; secondly, a credible punishment strategy to enforce the contribution in the first stage.

By knowing each other's utility functions, Fehr and Schmidt [1999]'s punishment strategy contains the following elements:

- A group of enforcers who have high βs and low αs. These enforcers are very similar to the cooperators in the game without punishment; they care sufficiently about advantageous inequality while not been too upset about getting less than those non-enforcers. Only players satisfy this property can make the punishment credible.
- The amount of punishment given the number of such enforcers. Since players' incentive to conduct punishment comes from their aversion towards inequality, the punishment would make the defector and enforcers' monetary payoffs equal.

Intuitively Fehr and Schmidt [1999]'s theory suggests that the incentive of an enforcer to conduct punishment comes from two part: one is to get rid of the disadvantageous inequality with the defectors in the first stage, the other is to avoid getting advantageous inequality with the other enforces. While the punishment is at players' own monetary cost as well as getting disadvantageous inequality with non-enforcers. However, in Bayesian games it is very tricky to find a proper punishment amount. Since everything is in expected terms, from an enforcers point of view, the number of enforcers among the other $n-2$ players (excluding the defector) can vary from 0 to $n-2$. Thus there exist situations of overly and insufficient punishment. If we simply making the defector's expected monetary payoff the same as the enforcers, then it's very difficult to analyze the effect of one enforcer's punishment on inequality (given other enforcers punish). We believe punishment need to be designed very carefully based on the number of players n.

While in a simple extension, by assuming n is sufficiently large that we can ignore the effect of inequality aversion towards a single defector. Then the model becomes exactly the same with the game without punishment. Enforcers can be defined in the same way as the cooperators in the previous sections. While by introducing the punishment scheme we are able to utilize some agent's inequality aversion to regulate others. The experimental results by Fehr et al. [1997] do show the striking change in the level of contribution once the punishment is allowed. Inspiring us to create other ways to utilize inequality aversion in incentive provision.

5.7 Conclusion and future work

In this chapter we did a preliminary study of linear public good games under incomplete game theory framework, where each players are not aware of the true preference parameters and can only infer other players actions based on a prior distribution of player's type. Unlike Fehr and Schmidt [1999]'s model, when players' preferences over inequality are privately known, the set of players that can be incentivized through aversion towards inequality will be further restricted.

We showed that in in public good game without punishment, we are able to characterize a group of cooperators who are willing to contribute among themselves. The contributors shares similar properties with Fehr and Schmidt [1999]'s model with complete information. However, there exist more than one way to define the contributors while Fehr and Schmidt [1999]'s theory only provides one typical case. Even if the condition in Fehr and Schmidt [1999]'s model doesn't hold, there might exist other possible equilibrium with positive contributions.

We also point out that Fehr and Schmidt [1999]'s comparison across the games are not very accurate. As in their public good game setting, the contribution vector is not revealed when there is no punishment. We showed theoretically how the equilibriums would differ when players compare the inequality on average base. Thus there is need to test whether revealing the contribution vector would yield different result in the public good game without punishment.

Throughout the theory we cannot rule out the existence of multi-equilibrium. Theory simply suggests players will cooperate, but why players are cooperating in some typical way remain unanswered. In this sense inequality aversion can help us when players have good faith upon others and it might not help at all under extreme circumstances. Apparently we would like to see what could be done in the later case to improve the efficiency. We can conduct possible linear public good experiments where players share are not equal, with one player's share greater than $\frac{1}{a}$. It would be interesting to see how other agents will react under these circumstances. On one hand, when one player's action can be perfectly predicted, other players would have information to predict other players contribution conditional on their types. On the other hand, introducing inequality in the sharing itself might bring disutility directly to the players. Thus it's important to check how players would react when different inequality measures are brought into the game.

Chapter 6

Conclusion

This dissertation analyses relational contracting in teams and the role inequality aversion can play in incentive provision. Through out our analysis, information plays an important role. How a piece of information can be utilized in incentive provision relies crucially on its quality, i.e. whether the information is publicly observable, does it contain any noise or is it at aggregate level or individual level. Ideally we all wish to have perfect verifiable public information signals at individual level so that we can easily solve the moral hazard problem. However, most real life challenges are more complicated than the ideal situation and we need to search for ways to mitigate the efficiency loss caused by information asymmetry.

We classify the information into two parts: the hard information, which is publicly observable and can be verified by outsiders; the soft information that can only be observed by agents who participate in production. If a piece of information can be verified by outsiders, a contract written contingent on this information can be enforced by court, therefore everyone has to obey what's been agreed. However, if something is only known within the production team itself, there is no way to enforce the agreements in a one shot game. Relational contract provides us a mechanism to utilize the "uncontractable" information in games that are played infinitely repeatedly. If an agent should pay some money to the rest of the team according to a non-court enforceable agreement and he regretted, then the agent breaks the relationship and gets a one-time benefit from reneging. The design of relational contract will need to make sure for all agents, maintaining the relationship is more valuable than any short-term benefits. Here is where the name "relational contract" comes from.

Throughout the dissertation we assume the output of team production is verifiable, while the information structure varies across different models. In chapter 3 we first looked into a model with deterministic production and stochastic individual performance signals. We find that an aggregate deterministic output is strong enough to conduct efficient punishment when the production game is played infinitely repeatedly. Intuitively, if the output is different to the output level under pre-agreed effort, then we can be 100% sure that shirking occurs. Then trigger strategy can be utilized to punish the whole group without further requirement of finding out who shirks. We showed that under this setting, relational contracting cannot do better than trigger strategy as the mechanism under these two kinds of incentive provision are exactly the same: the threat of breaking the relationship. Thus even when trigger strategy would not always be applicable, We further looked into non-deterministic production. We find out that the conditions posed by Rayo [2007] are not sufficient to ensure the validity of first-oder approach. However, we do justify the linear assumption made by Rayo [2007] when the first-order approach is valid. Further extension has been made to include negative sharing. We show that the optimal sharing structure is very different to the case with non-negative sharing. If negative sharing is allowed, then there will be one agent concentrating all the negative shares and make all other agents getting the full share. However, the nature of relational contract suggests that, if an agent receives no implicit incentive, then his/her surplus received would not affect incentive provision. While for the agent receiving negative shares, he needs to be motivated by large surplus from the production. We questioned ourselves what should be our definition of the endogenous principal, receiving the full share or getting the most surplus?

The previous question will be answered in chapter 4, where we pointed out a big gap not been observed by Rayo [2007]. In chapter 4 we studied a model with deterministic production and aggregate sub-team performances. We derived a simple and applicable grouping mechanism under which we only need to maintain some slight differences of sub-teams' organizational structure to provide strong relational incentives. We also present a simple one-step form of relational incentive and proved the rationality of it. The rest of the chapter focuses on the efficiency and sharing rules. We derived the necessary and sufficient condition for an outcome to be implemented with general sharing and relational contract. For general sharing we don't impose any restriction on the functional form of agents' shares, the sharing is specified point-wise.

The key contribution of the chapter 4 reveals the following property of linear sharing in relational contracts. When agent's surplus has endogenously become part of the incentive provision in relational contract, linear sharing would not give us the freedom to redistribute the team surplus. Thus Rayo [2007] cannot justify his maximization of team's surplus when assume linear sharing as there exist some outcomes that would be strictly preferred by some agents while not maximizing the team's surplus. In chapter 4 we conclude that if the social welfare function is the team's total surplus, then it's without loss to use linear sharing as linear sharing can enforce the efficient effort level under certain range of distributions of surplus. While linear sharing would not able to support extreme surplus distribution. Our result is further supported by a numerical example.

After a detailed analysis based on models where agents are purely incentivized through monetary payoffs, we turn our focus to inequality aversion in the fifth chapter. We admit there are much more work to be done in this area, the aim of our current research is to provide a view of inequality aversion models from Bayesian game perspective. We conclude that the equilibrium behavior in linear public good game hugely depends on how the inequality aversion is defined. We further points out the complementary experiments need to be done to Fehr and Schmidt [1999]'s model together with a suggestion of future study of our own interest.

Appendix A

Proofs of results in chapter 3

Proof for lemma 3.2

Proof. Suppose the optimal contract generate per period average surplus s^*. For an arbitrary optimal relational contract σ, suppose in each period t the contract implements $\mathbf{a}^t$ and generate surplus s^t, with $(1-\delta)\sum_{t=1}^{\infty}\delta^{t-1}s^t = s^*$. We first show the existence of a self-enforcing stationary contract that implements the initial effort schedule a^1 in each period generating per period average surplus as s^1, then we shall prove this contract is optimal.

Observe that s^1 must be at lease as large as s^*, because otherwise there would exist a history h^1 and a contract σ' that specifies the same action as σ, following history h^1 but one period in advance, creating a higher surplus than s^*

For each agent i and each period t, suppose σ specifies the sharing rule as s_i^t and the discretionary payment as p_i^t, where p_i^t can depend on all pubic history. Let

$\mu_i(l^1, y^1)$ denote the expected continuation payoffs achieved under σ from period 2 onwards.

In the first period, σ must satisfy the following constraints:

For each i

$$\begin{aligned} s_i^1(f(\mathbf{a}^1)) - c(a_i^1) + E[p^1(l^1, y^1)|\mathbf{a}^1] + \delta\mu_i(l^1, y^1), \\ \geq s_i^1(f(a_i', \mathbf{a}_{-i}^1)) - c(a_i') + E[p^1|a_i', \mathbf{a}_{-i}^1] \\ + \delta\mu_i(f_1^1(a_i', \mathbf{a}_{-i}^1), \ldots, f_m^1(a_i', \mathbf{a}_{-i}^1)) \end{aligned} \tag{A.1}$$

$$p^1(l^1, y^1) \geq \frac{\delta}{1-\delta}\bar{\pi}_i \tag{A.2}$$

Now define the new discretionary payment as $\hat{p} = p^1 + \delta\mu_i(l^1, y^1) - \delta\mu_i(l^1, y^1|\mathbf{a}^1))$. The idea is to move all the possible variation in the continuation payoffs into the discretionary payment in the current period. Let the new stationary contract $\hat{\sigma}$ propose sharing rule s_i^1 and discretionary payment $\hat{p}$ in each period. It's easy to verify the IC constraint remain the same, thus effort level $\mathbf{a}^1$ can be implemented while generating surplus s^1 in each period.

If the stationary contract $\hat{\sigma}$ is self-enforcing, we must have:

$$\hat{p} + \frac{\delta}{1-\delta}s^1 \geq \frac{\delta}{1-\delta}\bar{\pi}_i \tag{A.3}$$

Given $s^1 \geq s^*$, we would have

$$\frac{\delta}{1-\delta}s^1 \geq \delta\mu_i(l^1, y^1|\mathbf{a}^1) \tag{A.4}$$

(A.4) and (A.2) directly implies (A.3), thus $\hat{\sigma}$ is self-enforcing.

The existence of $\hat{\sigma}$ implies $s^1 \leq s^*$, because otherwise $\hat{\sigma}$ would generate a per period average surplus greater than σ which is already optimal. Since $s^1 \geq s^*$, we must have $s^1 = s^*$. Thus the stationary contract $\hat{\sigma}$ is optimal. □

Proof for theorem 3.3

Proof. For initiation of the production, team members' individual rationality constraints must be satisfied for any feasible sharing rule s:

$$s_i(y(\hat{a})) - c(\hat{a}_i) \geq \bar{\pi}_i \tag{A.5}$$

We will show that the contract is incentive compatible:

For any i, since $\hat{a}$ is the first-best effort level, $\hat{a}$ maximizes $y(a) - \sum c_i$. For all $a_i > \hat{a}_i$. The following inequality hold:

$$y(\hat{a}) - c_i(\hat{a}_i) - \sum c_{-i}(\hat{a}_{-i}) > y(a, \hat{a}_{-i}) - c_i(a) - \sum c_{-i}(\hat{a}_{-i})$$

which is equivalent to:

$$y(\hat{a}) - c_i(\hat{a}_i) > y(a, \hat{a}_{-i}) - c_i(a) \tag{A.6}$$

Thus team member's have no incentive to take effort level higher than $\hat{a}$.

If agent i renege and take effort level $a_i < \hat{a}_i$, since the output is increasing in a, $y(a_i, \hat{a}_{-i}) < y(\hat{a})$, the relationship will end from the beginning of the next period. To ensure the agent cannot get more under a_i, we require that $\pi_i(\hat{a}) + \frac{\delta}{1-\delta} pi_i(\hat{a}) \geq \pi_i(a_i, \hat{a}_{-i}) + \frac{\delta}{1-\delta}\bar{\pi}_i$, for sufficient large discount factor, the inequality should hold if for all individual, IR constraints hold with strict inequality, and this can be done by adjusting the constant payment in sharing rules(Levin [2003]). □

Proof for theorem 3.4

Proof. Suppose the optimal relational contract specifies the payment function for each agent i as (s_i, p_i). We construct a contract with trigger strategy as follows: For each agent i the share is specified by

$$\sim_i^t = \begin{cases} s_i(y^t) + E[p_i|a^*] & \text{if } y^t = y(\mathbf{a}^*) \\ s_i(y^t) & \text{if } y^t \neq y(\mathbf{a}^*) \end{cases}$$

, and trigger strategy is described by:

$$d_i^t = \begin{cases} 1 & \text{if } y^{t-1} \geq y(\mathbf{a}^*) \\ 0 & \text{if } y^{t-1} < y(\mathbf{a}^*) \end{cases}$$

if $y = y(a^*)$ then the production continues in the next period, if not then there will be no production in the next period and every agent receive their outside options. We then show the contract with trigger strategy is incentive compatible.

From the assumption that $\mathbf{a}^*$ can be implement by a relational contract, we have for each agent i, $s_i(y(a^*))+E[p_i|a^*] \geq s_i(y(a_i', a_{-i}^*))+E[p_i|a_i', a_{-i}^*] \geq s_i(y(a_i', a_{-i}^*))+ \inf p_i$,The self-enforcing constraint suggests, $\inf p_i \geq \frac{\delta}{1-\delta}(\bar{\pi} - \pi^*)$, so we have $s_i(y(a^*))+E[p_i|a^*] \geq s_i(y(a_i', a_{-i}^*))+\frac{\delta}{1-\delta}(\bar{\pi}-\pi^*)$, rearranging the equation we can show that the IC constraint for the trigger strategy contract holds: $s_i(y(a^*)) + E[p_i|a^*] + \frac{\delta}{1-\delta}\pi^* \geq s_i(y(a_i', a_{-i}^*)) + \frac{\delta}{1-\delta}\bar{\pi}$ □

Proof for theorem 3.6

Proof. We would first check that the first order condition hold with the new linear sharing rule:

For each i, the discretionary payment $g_i(l)$ remains the same. Thus reneging constraint will not be affected. Moreover It's easy to check that IC constraint is exactly the same on a^*.

We then show that the linear sharing rule is balanced. Since the original sharing rule balanced the budget, we have, for each y

$$\sum_{i=1}^{n} s_i(y) = y \tag{A.7}$$

From the linearity of expectations, we have

$$\sum_{i=1}^{n} E[s_i(y)|\mathbf{a}^*] = E[y|\mathbf{a}^*] \tag{A.8}$$

, for each i, let $x_i(E[y]) = E[s_i(y)]$, at the optimal point we have:

$$\frac{\partial E[s_i(y)|a^*]}{\partial a_i} = \frac{\partial x_i(E[y|a^*])}{\partial E[y]} \frac{\partial E[y|a^*]}{\partial a_i} \tag{A.9}$$

Thus we have $\alpha_i = \frac{\partial x_i(E[y|a^*])}{\partial E[y]}$, since $\sum_{i=1}^{n} x_i(E[y|\mathbf{a}^*]) = E[y|\mathbf{a}^*]$, then $\sum_{i=1}^{n} \alpha_i = \sum_{i=1}^{n} \frac{\partial x_i(E[y|a^*])}{\partial E[y]} = 1$.

By setting $\beta_i = (E[f_i(y)|a^*]) - f_i' E[y|a^*]E[y|a^*]$ we ensure that each agent receives the same expected surplus from the linear sharing rule from the original contract.

□

Proof for lemma 3.7

Proof. If all agent has the same marginal rate of substitution between explicit and implicit incentive, as the total shares among agents need to sum up to 1, the sum of implicit incentive among agents thus will be the same for any linear sharing rule. The sharing rule will have no impact on the self-enforcing constraint. Thus in this case, all linear sharing is optimal.

□

Proof for lemma 3.8

Proof. Since $\frac{\partial v_i}{\partial \alpha_i} = -(E[y|\mathbf{a}^*])'/\tau'$ is constant given $\mathbf{a}^*$,

$\frac{\partial h_i}{\partial \alpha_i} = (E[y|\mathbf{a}^*])'' - \frac{(E[y|\mathbf{a}^*])'}{\tau'}\tau''$ is invariant with respect to the share α_i □

Proof for theorem 3.9

Proof. If there are two agents i and j such that, $\frac{\partial v(\mathbf{a}^*)}{\partial \alpha_i} > \frac{\partial v(\mathbf{a}^*)}{\partial \alpha_j}$, it's always optimal to decrease α_i and raise α_j to the maximum, such that the sum of implicit incentive power of the two agents will be minimized. Thus for $n > 2$ agents, to minimize the implicit incentive needed, it's always optimal to decrease the share for agents with minimum marginal rate of substitution and raise the share for all other agents to 1. □

Appendix B

Proofs of results in chapter 4

Proof for lemma 4.2

Proof. Suppose the optimal contract generate per period average surplus s^*. For an arbitrary optimal relational contract σ, suppose in each period t the contract implements $\mathbf{a}^t$ and generate surplus s^t, with $(1-\delta)\sum_{t=1}^{\infty}\delta^{t-1}s^t = s^*$. We first show the existence of a self-enforcing stationary contract that implements the initial effort schedule a^1 in each period generating per period average surplus as s^1, then we shall prove this contract is optimal.

Observe that s^1 must be at lease as large as s^*, because otherwise there would exist a history h^1 and a contract σ' that specifies the same action as σ, following history h^1 but one period in advance, creating a higher surplus than s^*

For each agent i and each period t, suppose σ specifies the sharing rule as s_i^t and the discretionary payment as p_i^t, where p_i^t can depend on all pubic history.

Let $\mu_i(f_1^1(\mathbf{a}^1), \dots, f_m^1(\mathbf{a}^1))$ denote the continuation payoffs achieved under σ from period 2 onwards.

In the first period, σ must satisfy the following constraints:

For each i

$$\begin{aligned} s_i^1(\mathcal{F}(\mathbf{a}^1]) - c(a_i^1)) + p^1(f_1^1(\mathbf{a}^1), \dots, f_m^1(\mathbf{a}^1)) + \delta\mu_i(f_1^1(\mathbf{a}^1), \dots, f_m^1(\mathbf{a}^1)), \\ \geq s_i^1(\mathcal{F}(a_i', \mathbf{a}_{-i}^1)) - c(a_i') + p^1(f_1^1(a_i', \mathbf{a}_{-i}^1), \dots, f_m^1(a_i', \mathbf{a}_{-i}^1)), \\ + \delta\mu_i(f_1^1(a_i', \mathbf{a}_{-i}^1), \dots, f_m^1(a_i', \mathbf{a}_{-i}^1)) \end{aligned} \quad \text{(B.1)}$$

$$p^1(f_1, \dots, f_m) + \delta\mu_i(f_1, \dots, f_m) \geq \frac{\delta}{1-\delta}\bar{\pi}_i \quad \text{(B.2)}$$

Now define the new discretionary payment as $\hat{p}(f_1, \dots, f_m) = p^1(f_1, \dots, f_m) + \delta\mu_i(f_1, \dots, f_m) - \delta\mu_i(f_1^1(\mathbf{a}^1), \dots, f_m^1(\mathbf{a}^1))$. The idea is to move all the possible variation in the continuation payoffs into the discretionary payment in the current period. Let the new stationary contract $\hat{\sigma}$ propose sharing rule s_i^1 and discretionary payment $\hat{p}$ in each period. It's easy to verify the IC constraint remain the same, thus effort level $\mathbf{a}^1$ can be implemented while generating surplus s^1 in each period.

If the stationary contract $\hat{\sigma}$ is self-enforcing, we must have:

$$\hat{p}(f_1, \dots, f_m) + \frac{\delta}{1-\delta}s^1 \geq \frac{\delta}{1-\delta}\bar{\pi}_i \quad \text{(B.3)}$$

Given $s^1 \geq s^*$, we would have

$$\frac{\delta}{1-\delta}s^1 \geq \delta\mu_i(f_1^1(\mathbf{a}^1), \ldots, f_m^1(\mathbf{a}^1)) \tag{B.4}$$

(B.4) and (B.2) directly implies (B.3), thus $\hat{\sigma}$ is self-enforcing.

The existence of $\hat{\sigma}$ implies $s^1 \leq s^*$, because otherwise $\hat{\sigma}$ would generate a per period average surplus greater than σ which is already optimal. Since $s^1 \geq s^*$, we must have $s^1 = s^*$. Thus the stationary contract $\hat{\sigma}$ is optimal. □

Proof for theorem 4.5

Proof. The payment function 4.5 states the following:

If no one deviate in effort, then $\mathfrak{h}^{\mathbf{G}}(f_1, \ldots, f_m) = \emptyset$ and every agent's discretionary payment will be 0.

If deviation occurs then all agent who are detected by the monitoring scheme will receive a punishment $\frac{\delta}{1-\delta}(\bar{\pi}_i - \pi_i)$ which binds the self-enforcing constraint. All the other agents will then receive payment $w_i(\mathfrak{h})$ to balance the budget. Perfect monitoring scheme ensures two properties: 1. any agent who deviate will be receiving punishment. 2. For any unilateral deviation, there are always agents who will not be punished that can balance the budget.

With discretionary payment function (4.5), for any agent i, his IC constraint can be written as:

$$s_i(\mathcal{F}(a_i, \mathbf{a}_{-i})) - c_i(a_i) \geq s_i(\mathcal{F}(a_i', \mathbf{a}_{-i})) - c_i(a_i') + \frac{\delta}{1-\delta}(\bar{\pi}_i - \pi_i)$$

Rearranging the equation we have:

$$s_i(\mathcal{F}(a_i, \mathbf{a}_{-i})) - s_i(\mathcal{F}(a_i', \mathbf{a}_{-i})) - \frac{\delta}{1-\delta}(\bar{\pi}_i - \pi_i) \geq c(a_i) - c_i(a_i')$$

We denote the LHS of the above equation as incentive power. To prove (4.5) is optimal, we shall show that for any deviation in effort, there exist no other payment functions that creates higher incentive power.

For any agent i given s_i and $\mathbf{a}$ and p_i' that satisfies the self-enforcing constraint, the incentive power for any deviation:

$$\begin{aligned} & s_i(\mathcal{F}(a_i, \mathbf{a}_{-i})) - s_i(\mathcal{F}(a_i', \mathbf{a}_{-i})) - p_i'(a_i', \mathbf{a}_{-i}) \\ \leq & s_i(\mathcal{F}(a_i, \mathbf{a}_{-i})) - s_i(\mathcal{F}(a_i', \mathbf{a}_{-i})) - \inf p_i' \\ \leq & s_i(\mathcal{F}(a_i, \mathbf{a}_{-i})) - s_i(\mathcal{F}(a_i', \mathbf{a}_{-i})) - \frac{\delta}{1-\delta}(\bar{\pi}_i - \pi_i) \end{aligned}$$

, which completes the proof. □

Proof for lemma 4.8

Proof. The relation of α_i and a_i' can be described by the following first order condition: $\alpha_i \mathcal{F}'(a_i', a_-i) - c'(a_i') = 0$

Hence α_i can be expressed as a function of a_i', where $\alpha_i = \frac{c'(a_i')}{\mathcal{F}'(a_i', a_-i)}$. Since both $\mathcal{F}$ and c are twice differentiable on a_i', α_i is differentiable on a_i' and $\frac{d\alpha_i}{da_i'} =$

$\frac{c''(a'_i)f'(a'_i)-c'(a'_i)f''(a'_i)}{(\mathcal{F}'(a'_i))^2}$, given the strict convexity of the cost function $c(.)$ and the strict concavity of the production function $\mathcal{F}(.)$, $\frac{d\alpha_i}{da'_i} > 0$.

Thus the mapping $a'_i \to \alpha_i$ is a one to one differentiable function on $\mathbb{R}^+$, and we have the inverse mapping $\alpha_i \to a'_i$ is differentiable on $\mathbb{R}^+$ 1 ☐

Proof for lemma 4.9

Proof. Rearrange the IC constraint to make all terms related to α_i to be in the LHS, we have $\frac{1}{1-\delta}[s_i(\mathcal{F}(a_i, \mathbf{a}_{-1})) - c(a_i)] - [s_i(\mathcal{F}(a'_i, \mathbf{a}_{-1})) - c(a'_i)] \geq \frac{\delta}{1-\delta}\bar{\pi}_i$, where $a'_i \in \arg\max[s_i(\mathcal{F}(a'_i, \mathbf{a}_{-1})) - c(a'_i)]$. Based on lemma (4.8), we differentiate the LHS of the inequality with respect to α_i and have:

$$\begin{aligned}
&\mathcal{F}(\mathbf{a})/1-\delta-\mathcal{F}(a'_i, a_i) - [\frac{d\mathcal{F}(a'_i, a_i)}{d\alpha_i} - \frac{dc(a'_i)}{d\alpha_i}] \\
=&\mathcal{F}(\mathbf{a})/1-\delta-\mathcal{F}(a'_i, a_i) - [\frac{d\mathcal{F}(a'_i, a_i)}{da'_i} - \frac{dc(a'_i)}{da'_i}]\frac{da'_i}{d\alpha_i} \\
=&\mathcal{F}(\mathbf{a})/1-\delta-\mathcal{F}(a'_i, a_i) \\
>&0
\end{aligned} \tag{B.5}$$

Thus, the larger the α_i, the larger $\frac{1}{1-\delta}[s_i(\mathcal{F}(a_i, \mathbf{a}_{-1}))-c(a_i)]-[s_i(\mathcal{F}(a'_i, \mathbf{a}_{-1}))-c(a'_i)]$ and thus the inequality would be more likely to hold. ☐

Proof for lemma 4.7

Proof. Necessity: It's directly shown by summing up all agents' IC constraints together.

Sufficiency: If inequality (4.13) holds but for some agents, the IC constraints do not hold, it's always possible to transfer some fixed payment β from other agents whose IC constraints hold with strict inequality such that inequality (4.13) will not be influenced and all IC constrains will hold. □

Proof for theorem 4.10

Proof. Kuhn-Tucker first order condition for problem (4.15) suggest that: for all i whose deviation effort $a'_i > 0$, we have:

$$\mathcal{F}(a'_i, a_{-i}) + \alpha_i \frac{\partial \mathcal{F}(a'_i, a_{-i})}{\partial a_i} * \frac{\partial a_i}{\partial \alpha_i} - \frac{\partial c(a'_i)}{\partial a_i} * \frac{\partial a_i}{\partial \alpha_i} + \lambda = 0$$

, where λ is the Lagrangeun mutiplier for budget balancing constraint. For all $a'_i > 0$ we also have: $\alpha_i \frac{\partial \mathcal{F}(a'_i, a_{-i})}{\partial a_i} - \frac{\partial c(a'_i)}{\partial a_i} = 0$, which directly implies $\mathcal{F}(a'_i, a_{-i}) = -\lambda$.

Suppose s^* is optimal, if with s^* every agent's deviation effort $a'_i > 0$, then theorem holds from the Kuhn-Tucker first order condition. If for some agents $a'_i = 0$, we need to prove for each of these agents, $\mathcal{F}(a'_i, a_{-i})$ must be the same as everyone else.

Let $\mathcal{F}^*$ denote the deviation output level for all agents with $a'_i > 0$. For an arbitrary agent j, with $a'_j = 0$;

If $\mathcal{F}(0, a_{-i}) > \mathcal{F}^*$, a marginal decrease in α_j would decrease the objective function by $\mathcal{F}(0, a_{-i})$, while for all other agents whose target effort level is positive, a same marginal increase in α_i would increase the objective function by $\mathcal{F}^*$, thus the value objective function can be further reduced by shifting such agents' share down.

If $\mathcal{F}(0, a_{-i}) < \mathcal{F}^*$, for the same reasoning the objective function is not at the minimum level since raising agent j's share without breaking the budget constraint will reduce the value of the objective function.

Hence for any agent j whose $a'_j = 0$, we must have $\mathcal{F}(0, a_{-i}) = \mathcal{F}^*$ $\square$

Proof for lemma 4.11

Proof. Under linear sharing, if each agent i's optimal deviation output is $\hat{\theta}$, then $\alpha_i = v'_i(\hat{\theta})$ and $\sum_{i=1}^{n} alpha_i = 1$.

With non-linear sharing, if each agent deviates to the same output level θ'. Since $\sum_{i=1}^{n} s_i(\theta') = \theta'$, $\sum_{i=1}^{n} s'_i(\theta') = 1$. If $\theta' \neq \hat{\theta}$, then there exist another balanced linear sharing such that $\alpha'_i = s'_i(\theta')$ making every agent deviate to $\theta\prime$. Since, $v'_i(\theta)$ is strictly increasing with θ for each i, $\sum_{i=1}^{n} \alpha'_i \neq 1$ which contradict to the previous condition. Thus $\theta' = \hat{\theta}$. $\square$

Proof for lemma 4.13

Proof. Since $\theta^* \in \arg\max\{\theta - \sum_{i=1}^{n} v_i(\theta)\}$, the following FOC must hold:

$$1 - \sum_{i=1}^{n} \frac{dv_i(\theta^*)}{d\theta} = 0 \tag{B.6}$$

For each i, let $a_i(\theta^*) = a_i^*$, we have:

$$\frac{dv_i(\theta^*)}{d\theta} = \frac{dc_i(a_i(\theta^*))}{da} * \frac{da_i(\theta^*)}{d\theta} = \frac{c_i'(a^*)}{\theta'(a_i^*)}$$

We construct a linear sharing such that $\alpha_i = \frac{dv_i(\theta^*)}{d\theta}$, it's clear the sharing is balanced such that $\sum_{i=1}^{n} \alpha_i = 1$. We can also show that with α_i, each agent i's optimal deviation will be θ^*.

With α_i, each agent solves for $\alpha_i\theta(a_i) - c_i(a_i)$ and the FOC: $\alpha_i\theta'(a) - c_i'(a_i) = 0$ holds on a_i^*.

From lemma 4.11, when each agent i's optimal deviation level is the same, then this deviation output is uniquely defined as $\hat{\theta}$. So we have $\hat{\theta} = \theta^*$ □

Proof for theorem 4.14

Proof. If $\theta^* - \sum_{i=1}^{n} v_i(\theta^*) \leq \frac{1}{1-\delta}[\mathcal{F}^* - \sum_{i=1}^{n} c_i(a^*)] - \frac{\delta}{1-\delta}\sum_{i=1}^{n} \bar{\pi}_i$.

Define α_i as what we did in the proof of lemma 4.13. We can ensure θ^* maximize $\alpha_i\theta - v_i(\theta)$ for each i.

Given $\beta = (\beta_i, \ldots, \beta_n)$, with $\sum_{i=1}^{n} \beta_i = 0$, if

$$\alpha_i\theta^* + \beta_i - c_i(a_i) \geq \frac{1}{1-\delta}[\alpha_i\mathcal{F}^* + \beta_i - c_i(a^*)] - \frac{\delta}{1-\delta}\bar{\pi}_i$$

Define D_i as the difference of LHS and RHS on the above inequality, we can write $D_i = -\frac{\delta}{1-\delta}\beta_i + C_i$, where C_i is the constant

$$\alpha_i\theta^* - c_i(a_i) - \frac{1}{1-\delta}[\alpha_i\mathcal{F}^* - c_i(a^*)] + \frac{\delta}{1-\delta}\bar{\pi}_i$$

Since $0 < \delta < 1$, D_i is strictly decreasing with β_i for all i with the same marginal rate across all agent i. For all i with $D_i > 0$, take $\triangle\beta_i = \frac{1-\delta}{\delta}D_i$, we alter β_i such that with $\beta_i' = \beta_i + \triangle\beta_i$, we have $D_i' = 0$

If we do the same with for each j of the rest of the agents with $D_j \leq 0$, we must have $|\sum\triangle\beta_j| \geq |\sum\triangle\beta_i|$, suggesting that if we transfer $-\sum\triangle\beta_i$ to agents with $D_j \leq 0$, we can ensure $\sum D_j' \leq 0$ while $D_i' = 0$ for all i.

If any $D_j' > 0$, repeat the above steps we can ensure $D_i \leq 0$ for all agents. $\square$

Appendix C

Proofs of results in chapter 5

Proof for theorem 5.1

Proof. Without loss of generality we label i as 1, consider an arbitrary contribution vector $(g_2, \ldots, g_n)$ of other players. Suppose $\beta_1 < 1 - \frac{a}{n}$, if player 1 contribute 0, his utility function is given as:

$$U_1(g_1 = 0) = y + \frac{a}{n}\sum_{i=2}^{n} g_i - \frac{\beta_i}{n-1}\sum_{i=2}^{n} g_i \tag{C.1}$$

If player 1 contribute $g_1 > 0$, we could have the following three situations:

1. If $g_1 \leq \min\{g_2, \ldots, g_n\}$, we have

$$U_1(g_1 > 0) = y + \frac{a}{n}\sum_{i=2}^{n} g_i + (\frac{a}{n} - 1)g_1 - \frac{\beta_1}{n-1}\sum_{i=2}^{n}(g_i - g1) \tag{C.2}$$

, rearranging the terms we have: $U_1(g_1 > 0) = U_1(g_1 = 0) + (\frac{a}{n} - 1 + \beta_1)g_1$. Since $\beta_1 < 1 - \frac{a}{n}$, we have $U_1(g_1 > 0) < U_1(g_1 = 0)$

2. If $\min\{g_2, \ldots, g_n\} < g_1 \leq \max\{g_2, \ldots, g_n\}$, let $M = \{i|g_i < g_1\}$ and $N = \{j|g_j \geq g_1, j > 1\}$ and let k denote the number of elements in M, we have

$$U_1(g_1 > 0) = y + \frac{a}{n}\sum_{i=2}^{n} g_i + (\frac{a}{n} - 1)g_1 - \frac{\alpha_1}{n-1}\sum_{i \in M}(g_1 - g_i) - \frac{\beta_1}{n-1}\sum_{j \in N}(g_j - g1)$$

If $\alpha_1 \geq \beta_1 \geq 0$, then

$$\begin{aligned} U_1(g_1 > 0) \leq & y + \frac{a}{n}\sum_{i=2}^{n} g_i + (\frac{a}{n} - 1)g_1 - \frac{\beta_1}{n-1}\sum_{i \in M}(g_1 - g_i) - \frac{\beta_1}{n-1}\sum_{j \in N}(g_j - g1) \\ \leq & y + \frac{a}{n}\sum_{i=2}^{n} g_i + (\frac{a}{n} - 1)g_1 + \frac{\beta_1}{n-1}\sum_{i \in M}(g_1 - g_i) - \frac{\beta_1}{n-1}\sum_{j \in N}(g_j - g1) \\ = & y + \frac{a}{n}\sum_{i=2}^{n} g_i + (\frac{a}{n} - 1)g_1 - \frac{\beta_1}{n-1}\sum_{i=2}^{n}(g_i - g1) \\ = & U_1(g_1 = 0) + (\frac{a}{n} - 1 + \beta_1)g_1 \end{aligned}$$

$$< U_1(g_1 = 0)$$

If $\alpha_1 \geq 0 > \beta_1$, then it's straight forward that the agent 1's payoff is strictly decreasing with his contribution g_1 as contributing will decrease his advantageous inequality and increase his disadvantageous inequality with other agents.

3. If $g_1 > \max\{g_2, \ldots, g_n\}$, we have

$$\begin{aligned}
U_1(g_1 > 0) <& y + \frac{a}{n}\sum_{i=2}^{n} g_i + (\frac{a}{n} - 1)g_1 - \frac{\alpha_1}{n-1}\sum_{i=2}^{n}(g_1 - g_i) \\
<& y + \frac{a}{n}\sum_{i=2}^{n} g_i + (\frac{a}{n} - 1)g_1 + \frac{\beta_1}{n-1}\sum_{i=2}^{n}(g_1 - g_i) \\
=& U_1(g_1 = 0) + (\frac{a}{n} - 1 + \beta_1)g_1 \\
<& U_1(g_1 = 0)
\end{aligned}$$

Therefore, if $\beta_1 < 1 - \frac{a}{n}$, $U_1(g_1 > 0) < U_1(g_1 = 0)$ for any $(g_2, \ldots, g_n)$. □

Proof for theorem 5.3

Proof. If $\max\{\int_x^y \int_\beta^y g(\alpha|\beta)f(\beta)d\alpha d\beta - \frac{y+1-a/n}{y+x}\} < 0$, then there doesn't exist any interval $[x, y]$ such that $P(x \le \beta \le \alpha \le y) \ge \frac{y+1-a/n}{y+x}$, then there doesn't exist a strategy that can support positive contributions for certain types of players.

We have shown that condition 5.7 must hold if any player with type $\theta_i \in \Theta$'s best respond is to contributing g, if other players with $\theta_j \in \Theta$ are contributing g. Here we would show that the inequality must hold with equality.

Given x and y, if $P(x \le \beta \le \alpha \le y) \ge \frac{y+1-a/n}{y+x}$ hold with strict inequality, we would be able to find some players such that $\beta^* = x$ while $\alpha^* > y$ such that $P(x \le \beta \le \alpha \le y) \ge \frac{\alpha^*+1-a/n}{\alpha^*+\beta^*} > \frac{y+1-a/n}{y+x}$. Since $P(x \le \beta \le \alpha \le y)$ is increasing in y and decreasing in x, $P(\beta^* \le \beta \le \alpha \le \alpha^*) \ge P(x \le \beta \le \alpha \le y)$. Suggesting player with $[\alpha^*; \beta^*]$ should have incentive to contribute g as well which contradict to the fact that $[\alpha^*; \beta^*] \notin \Theta$ □

Proof for theorem 5.4

Proof. Sufficiency: If $\beta_i > 1 - \frac{a}{n}$ for all i, then for any symmetric contribution vector with $g_i = g \in [0, y]$ can be supported as a Bayesian Nash equilibrium. For any player i, if all the other players contribute g, we have $\bar{g}_i = g$. Since $\beta_i \geq 1 - \frac{a}{n}$, player i the marginal benefit from contributing less is always less or equal to the marginal disutility from inequality aversion. While any contribution higher than g would give the player strictly less utility as contribution is costly and players are averse to disadvantageous inequality.

Necessity: If the following strategy: For each player: if player i's type $\theta_i \in \Theta$, then $g_i = g > 0$, otherwise $g_i = 0$ forms a perfect Bayesian Nash equilibrium. Then the expected payoff when player i's type $\theta_i \in \Theta$ and contributing g must be greater or equal to the expected payoff when player i's type $\theta_i \in \Theta$ and contributing other g'. Under this strategy, we have $E[\bar{g}_i] = p(\theta_i \in \Theta)g$.

Since $\alpha_i \geq \beta_i > 0$, g shall never be greater than $E[\bar{g}_i]$. On the other hand $E[\bar{g}_i] = p(\theta_i \in \Theta)g \leq g$. If g can be supported in the equilibrium, then $g = p(\theta_i \in \Theta)g$. □

Bibliography

Baker, G., Gibbons, R. and Murphy, K. J. [1994], 'Subjective performance measures in optimal incentive contracts', *The Quarterly Journal of Economics* **109(4)**, 1125–1156.

Bolton, G. E. and Ockenfels, A. [2000], 'A theory of equity, reciprocity, and competition', *American Economic Review* **90**, 166–193.

Carver, T. and Vondra, A. [1994], 'Alternative dispute resolution: Why it doesn't work and why it does', *Harvard Bussiness Review* **72**, 3–120.

Che, Y. and Yoo, S. [2001], 'Optimal incentives for teams', *American Economic Review* **91**, 525–541.

Engelmann, D. and Strobel, M. [2004], 'Inequality aversion, efficiency, and maximin preferences in simple distribution experiments', *American Economic Review* **94**, 857–869.

Eswaran, M. and Kotwal, A. [1984], 'The moral hazard of budget-breaking', *The RAND Journal of Economics* **15(4)**, 578–581.

Fehr, E., Gachter, S. and Kirchsteiger, G. [1997], 'Reciprocity as a contract enforcement device: Experimental evidence', *Econometrica* **65**(4), 833–860.

Fehr, E., Kirchsteiger, G. and Riedl, A. [1993], 'Does fairness prevent market clearing? an experimental investigation', *The Quarterly Journal of Economics* **108**(2), 437–459.

Fehr, E. and Schmidt, K. M. [1999], 'A theory of fairness competition and cooperation', *The Quarterly Journal of Economics* **114(3)**, 817–868.

Forsythe, R., Horowitz, J. L., Savin, N. and Sefton, M. [1994], 'Fairness in simple bargaining experiments', *Games and Economic Behavior* **6**(3), 347 – 369.

Green, J. R. and Stockey, N. L. [1983], 'A comparison of tournaments and contracts', *Journal of Political Economy* **913**, 349–364.

Holmstrom, B. [1982], 'Moral hazard in teams', *Bell Journal of Economics* **13**, 324–340.

Huang, P. and Wu, H. [1994], 'More order without more law: A theory of social norms and organizational cultures', *Journal of Law, Economics and Organization* **10**, 390–406.

Itoh, H. [1991], 'Incentives to help in multi-agent situations', *Econometrica* **59**, 611–636.

Itoh, H. [2004], 'Moral hazard and other-regarding preferences', *Japanese Economic Review* **55**, 18–45.

Kambe, S. [2006], 'Subjective evaluation in agency contracts', *Japanese Economic Review* **57**, 121–140.

Kim, S. K. and Wang, S. [1998], 'Linear contracts and the double moral-hazard', *Journal of Economic Theory* **82**(2), 342–378.

Kvaloy, O. and Olsen, T. [2006], 'Team incentives in relational employment contracts', *Journal of Labor Economics* **24**, 139–170.

Levin, J. [2003], 'Relational incentive contracts', *American Economic Review* **93**, 835–857.

MacLeod, W. B. [2003], 'Optimal contracting with subjective evaluation', *The American Economic Review* **93**(1), 216–240.

Maestri, R. [2012], 'Bonus payments versus efficiency wages in the repeated principal-agent model with subjective evaluations', *American Economic Journal: Microeconomics* **4**, 34–56.

McAfee, R. P. and McMillan, J. [1991], 'Optimal contracts for teams', *International Economic Review* **32**(3), 561–577.

Mirrelees, J. [1976], 'The optimal structure of authority and incentives within an organization', *Bell Journal of Economics* **7**, 105–131.

Mookherjee, D. [1984], 'Optimal incentive schemes with many agents', *Review of Economic Studies* **51**, 433–446.

Nandeibam, S. [2002], 'Sharing Rule in Teams', *Journal of Economic Theory* **107**(3), 407–420.

Nandeibam, S. [2003], 'Implementation in Teams', *Economic Theory* **22**(3), 569–581.

Plambeck, E. L. and Taylor, T. A. [2006], 'Partnership in a Dynamic Production System with Unobservable Actions and Noncontractible Output', *Management Science* **52**(10), 1509–1527.

Rabin, M. [1993], 'Incorporating fairness into game theory and economics', *American Economic Review* **83**, 1281–1302.

Rabin, M. [2002], 'A perspective on psychology and economics', *European Economic Review* **46**, 657–685.

Rayo, L. [2005], Relational team incentives and ownership, Job market papers.

Rayo, L. [2007], 'Relational incentives and moral hazard in teams', *The Review of Economic Studies* **74**, 937–963.

Rogerson, W. [1985], 'The first-order approach to principal-agent problems', *Econometrica* **53**, 1357–1368.

Winter, E. [2009], 'Incentive reversal', *American Economic Journal: Microeconomics* **1**(2), 133–147.