Coordination, Rent-Seeking and Control*

Sandeep Baliga
Northwestern University

Tomas Sjöström
Rutgers University

April 11, 2013

Abstract

Two agents want to coordinate their decisions but may also try to extract rents from each other. Decsions are negotiated at the interim stage, when the agents have private information. The agent who owns an asset has the right to make a unilateral decision regarding this asset, so negotiations must satisfy a participation constraint. We compare nonintegration, where each agent owns one asset, with integration, where one agent owns both assets. We derive simple necessary and sufficient conditions for the first best to be implemented with integration or nonintegration. These conditions can never be satisfied simultaneously, so integration sometimes dominates nonintegration and vice versa.

1 Introduction

In a world of complete contracts, outcomes would be independent of organizational form. Therefore, to understand why one organizational form might dominate another, contracts must be assumed to be incomplete (Williamson [23], [26]). Grossman and Hart [6] provided a theory of surplus maximizing control rights based on complete information and noncontractible ex ante investments. In our model, there are no ex ante investments. Instead, agents have private information about the state of the world. A complete contract

*We thank Robert Gibbons, Oliver Hart, Niko Matouschek and Andrew Newman for valuable comments.

would specify the decisions in all states, but we assume such complete contracting is impossible. In reality, the set of possible states would be extremely large, and it is unlikely that they could all be specified ex ante. But even without a complete contract, it may be possible to arrive at first best (surplus maximizing) decisions. Decisions are made at the interim stage, when the agents have private information, and are subject to interim participation constraints. But we make no further incomplete contracting assumption: any incentive-compatible agreement that satisfies the interim participation constraint is allowed. The participation constraints depend on control rights, i.e., on who has the right to make a decision if negotiations fail. We obtain a theory of surplus maximizing control rights based on private information and interim participation constraints.

For a concrete example, suppose two software companies can cooperate to create programs that work well together, say a word processing program and a spreadsheet program. The decisions to be made might concern what features to add to each program. Say each firm can either make its software program specialized so it largely performs its core task, or generalized by incorporating features of the other firm's program into its own. The optimal decisions depend on the state of the world. Any feasible decision making procedure must induce voluntary participation, given the firms private information about the state.

More formally, there are two agents A and B, and two decisions to be made, q_A and q_B. In the example, each software company would have private information about its cost of making a specific software program. In our model, each agent $i \in \{A, B\}$ has a privately known idiosyncratic preference over q_i. This preference is his *type* and is denoted θ_i. The state of the world is (θ_A, θ_B). However, decision q_i matters also to agent $j \neq i$.[1] To keep the model simple, the effect q_i has on agent j is common knowledge. For example, if the word processing program incorporates rudimentary spreadsheet capabilities, it is known that he profit of the spreadsheet program maker will be reduced.

The decisions q_A and q_B are negotiated at the interim stage, when the agents privately know their own types. To find out if the first best (surplus-maximizing) decisions can, in principle, be negotiated at the interim stage, we apply the revelation principle. An unbiased mediator proposes an incentive-compatible revelation mechanism for making decisions. This is not meant to

[1] Hart and Holmström [8], Alonso, Dessein and Matouschek [1] and Rantakari [20] study similar coordination problems.

be descriptive of real-world organizations, where agents presumably use some more intuitive bargaining procedure (not necessarily involving announcements of types). However, by using the revelation principle, we establish if surplus losses are necessary consequences of the inability to write complete contracts. Moreover, using a revelation mechanism at the interim stage does not contradict the inability to write complete contracts ex ante. Ex ante, the set of possible states may be very large and not describable, but interim, the set of possible types is known to lie in a much smaller set which we identify with $[0, 1]$. The incomplete contracts assumption is that it is impossible to describe the large set of states ex ante, but this does not preclude describing the small set of types which are known to be possible at the interim stage.

Interim negotiations must satisfy interim individual rationality (or participation) constraints which depend on the control rights, i.e., on who has the right to choose q_A and q_B. With *nonintegration*, agent A has the right to choose q_A and agent B has the right to choose q_B. With *integration*, one agent ("the boss") has the right to choose both q_A and q_B. Following Grossman and Hart [6], we may interpret the nonintegrated case as separate ownership of two assets A and B. With integration, one agent owns both assets. We find that the *first-best* surplus level can sometimes be implemented with integration but not with non-integration, and vice versa, depending on certain parameters. Thus, the (unique) ex ante surplus-maximizing allocation of control rights may be either integration or nonintegration, depending on the parameters.

We assume payoff functions depend on two parameters, ρ and γ. Parameter ρ indicates how much rent an agent can extract by behaving opportunistically (e.g., by incorporating features of the other firm's software program into its own). Parameter γ indicates the gain from coordination. We characterize parameter regions for which the first best can be implemented with integration or nonintegration, and find that these regions are non-overlapping. Thus, integration sometimes dominates nonintegration and vice versa. For parameter regions where achieving coordination is more important than preventing rent-seeking, the first best can be implemented with nonintegration but not with integration. Conversely, there are parameter regions where rent-seeking is a more important problem and the first best can be implemented with integration but not with nonintegration.

Specifically, we show that there is ρ^* (which depends on γ) such that the first best can be implemented with nonintegration if and only if $\rho \leq \rho^*$. To prove this result, we first identify the type of either agent who has most to gain from rent-seeking. This is the type whose IR constraint is most difficult to satisfy (since the first best revelation mechanism eliminates rent-seeking). Then we show that this type's reservation payoff is increasing in ρ. Thus, to satisfy IR, this type must get a higher expected payoff, the higher is ρ. But then the IC constraints imply that *all* types must get a higher expected payoff the higher is ρ. This holds for both agents, since they are symmetric under nonintegration. But the total surplus cannot be raised above the first best level. Therefore, there is ρ^* such that if $\rho > \rho^*$ then there is not enough surplus to satisfy all IC and IR constraints. This argument shows that if the incentives for opportunistic rent-seeking are too strong, nonintegration cannot implement the first best.

To understand if integration can do better, two cases are considered separately. Intuitively, if ρ is small compared to γ (the *cooperative case*) then the most important problem is to ensure coordination, but if ρ is large compared to γ (the *rent-seeking case*) then preventing rent-seeking is more important. Consider first the rent-seeking case. The argument of the previous paragraph shows why, in the rent-seeking case, the first best cannot be implemented with nonintegration. With integration, a more positive result is obtained. On the one hand, the boss's reservation payoff is high if ρ is high; if negotiations break down he will use his control rights to "hold up" (extract rents from) his subordinate and achieve a payoff of at least ρ. But on the other hand, the subordinate's reservation payoff is low if ρ is high, because he expects to be held up. Moreover, in our model rent-extraction is inefficient since it involves giving up coordination benefits, so as ρ increases the subordinate's reservation payoff decreases faster than the boss's reservation payoff increases. What matters for implementability of the first best is the *sum* of the reservation payoffs, so as ρ increases implementation becomes easier for the integrated firm. Intuitively, in the rent-seeking case the subordinate has a lot to lose if the boss is unconstrained, and therefore the subordinate is highly motivated to negotiate a first best agreement. As a result, the integrated firm can implement the first best if ρ is big (compared to γ).[2]

Finally, consider integration in the cooperative case. In this case, if negotiations break down the boss prefers to coordinate rather than hold up the subordinate agent (he will give up ρ in favor of γ). The subordinate has a

[2] This is reminiscent of Williamson's [25] idea that an exchange of hostages yields a threat point that can enforce efficient trade, and that in a repeated game, the one-shot equilibrium is a threat point which allows cooperation to take place in equilibrium (Halonen [7]).

type, type 0, who derives no idiosyncratic benefit from any decision. If negotiations break down, type 0 knows the outcome will be first best, because the boss will ensure coordination on the decision the boss prefers (and type 0 doesn't care about which decision they coordinate on). If type 0 is required by a first best revelation mechanism to make a monetary transfer to the boss, he will refuse to participate. Type 0 must get a high expected payoff from the mechanism – this is the binding IR constraint for the subordinate. But then, the IC constraints imply that *all* types of the subordinate must get high expected payoffs. Even those types of subordinate who, in principle, would be willing to pay a lot to implement the first best can always pretend to be indifferent (i.e., to be type 0). This threat of "haggling" implies that the subordinate must get large information rents. Naturally, the boss requires a high expected payoff to participate, since he has all the control rights. So, in the cooperative case, the boss and the subordinate both demand very high expected payoffs, but this violates budget balance. In the cooperative case the first best does not add enough surplus, compared to the disagreement point, to satisfy all IC and IR constraints for both agents. As a result, the integrated firm cannot implement the first best if ρ is small (compared to γ).

Williamson [23] argued that disagreements between independent (nonintegrated) agents may lead to costly "haggling". In our model, this haggling consists of strategic manipulation of private information, which may cause decisions to differ from the first-best (surplus maximizing) outcome. Williamson [23] further argued that disputes within an integrated firm can be settled more efficiently, because the boss can make unilateral decisions. However, to make an efficient decision, the boss would still need to get information from the subordinate. But as in the nonintegrated case, strategic manipulation of private information may then cause the decision to differ from the first-best. We use the revelation principle to compare the (unavoidable) surplus losses in the integrated firm with the (unavoidable) surplus losses in the nonintegrated case.

Milgrom and Roberts [16] provided an informal theory based on a comparison of the "bargaining costs" of nonintegration with the "influence costs" of integration. Influence costs occur when the subordinate expends time and effort on "influence activities". Milgrom and Roberts [16] emphasized the distinction between these two kinds of costs. In contrast, we treat nonintegration and integration symmetrically: in each case, decision making requires the revelation of private information, but strategic manipulation may prevent the first best from being implemented.[3] Baker, Gibbons and Murphy [2] argue that decision rights within an organization are not contractible: the boss cannot formally delegate decision making to an employee, because the boss always has the right to overturn the subordinate's decision. More generally, they argue that formal authority can only be allocated via asset ownership, although informal authority can be allocated in a repeated game. In equilibrium the agent who owns the asset must be better off than he would be by making a unilateral decision, since the latter option cannot be contracted away. This is similar to our participation constraint. But our underlying model is different from both Milgrom and Roberts [16] and Baker, Gibbons and Murphy [2], and unlike them we use a mechanism design approach at the interim stage.

Laffont and Maskin [12] and Myerson and Satterthwaite [17] showed that IC and interim IR constraints may preclude surplus maximization in public good and bilateral trading problems (where the decision is how much or whether to produce or trade and with what transfers). In a Myerson-Satterthwaite type trading problem, it is known that the attainable surplus can depend on property rights (Cramton, Gibbons and Klemperer [4], Matouschek [15]). Our underlying model is different: a decision has to be made for each asset, which raises issues of coordination and rent-extraction that do not exist in the Myerson-Satterthwaite framework, and leads to a richer analysis of the optimal allocation of control rights. Formally, our arguments closely follow Myerson and Satterthwaite [17].[4]

Section 2 contains the basic model and Section 3 describes the first best. In Section 4 we introduce incentive compatibility and individual rationality constraints and prove certain preliminary results. Section 5 contains our main result for nonintegration. In Sections 6 and 7 we consider integration in the cooperative and rent-seeking cases, respectively. Section 8 summarizes our main findings.

[3] The resulting surplus loss could of course be labelled a bargaining cost in one case and influence cost in the other.

[4] An alternative approach would rely on VCG mechanisms along the lines of Krishna and Perry [11], Makowski and Mezzetti [13], Neeman [18] and Williams [21].

2 Model

2.1 Decisions and Payoff Functions

Consider a relationship involving two agents, A and B. Two decisions have to be made, q_A and q_B. For simplicity, there are only two options, *Down* or *Up*, for each decision. For $i \in \{A, B\}$, let θ_i denote agent i's *type.* Agent i's type is his private information. Agent i's payoff function is $v_i(q, \theta_i) + t_i$, where $v_i(q, \theta_i)$ is the benefit type θ_i derives from decision profile $q = (q_A, q_B)$ and t_i is a monetary transfer. An *outcome* of the game consists of a decision profile $q = (q_A, q_B)$ and a transfer profile $t = (t_A, t_B)$. Budget balance requires that the transfers always sum to zero: $t_A + t_B = 0$. There are no individual limited liability constraints.

We represent $v_i(q, \theta_i)$ in the matrix (1). The row indicates q_i and the column indicates q_j.

$$\begin{array}{ccc} & Up & Down \\ Up & \theta_i + \gamma & \theta_i + \rho \\ Down & -\rho & \gamma \end{array} \tag{1}$$

The payoff matrix is highly stylized to simplify calculations, but it captures the basic idea that the relationship has cooperative, opportunistic (rent-seeking) and idiosyncratic elements. On the one hand, if the decisions are coordinated (either both *Up* or both *Down*) then $\gamma > 0$ is added to each agent's payoff. On the other hand, if $q_i =$ *Up* and $q_j =$ *Down* then agent i gains $\rho > 0$ at the expense of agent j.[5] Finally, θ_i (which can be positive or negative) represents an idiosyncratic benefit agent i gets from $q_i =$ *Up*.

There are two salient cases. In the *cooperative case* we have $\gamma > \rho$. In this case, the benefit of coordination dominates the possible gains from rent seeking, and the biggest number in payoff matrix (1) is always on the main diagonal. In the *rent-seeking case* we have $\rho > \gamma$. In this case, the gain from rent seeking dominates the benefit of coordination. The incentives to cooperate or rent-seek will determine the impact of integration on surplus maximization.

[5] In the example from the Introduction, *Up* corresponds to producing generalized software and *Down* to producing specialized software. A more general payoff structure, where the gains and losses from opportunism do not cancel out, would complicate the calculations without adding any new insights.

Types are independently drawn from a distribution with a differentiable c.d.f. F with support $\left[\underline{\theta}, \overline{\theta}\right]$. We make the following two assumptions on F.

Assumption 1. $F'(\theta_i) < 1/(2\gamma)$ for all $\theta_i \in \left[\underline{\theta}, \overline{\theta}\right]$.
Assumption 2. $\underline{\theta} = -\overline{\theta}$ and $F(\theta_i) = 1 - F(-\theta_i)$ for all $\theta_i \in \left[\underline{\theta}, \overline{\theta}\right]$.

Assumption 1 says that there is "sufficient uncertainty", that is, the density F' is rather flat and not concentrated around one point. This will be used to show uniqueness of equilibrium when each agent i chooses q_i noncooperatively. Assumption 2 says that F is symmetric around 0. Thus, the idiosyncratic element θ_i is equally likely to favor *Up* or *Down*. Finally, it will simplify the exposition to assume $2\gamma < \overline{\theta} < \infty$ but this has no substantive implications.

2.2 Time Line

The game has four stages.

Stage 0. The agents allocate control rights.

Stage 1. Each agent $i \in \{A, B\}$ privately observes his own type θ_i.

Stage 2. An impartial mediator proposes a mechanism Γ. Each agent (simultaneously) either accepts or rejects Γ. If both accept, then move to stage 3a, otherwise move to stage 3b.

Stage 3. (a) If both agents agreed to participate in Γ at stage 2, then messages are exchanged and an outcome is implemented as specified by Γ. (b) If at least one agent refused at stage 2, then under *integration* agent A chooses both q_A and q_B, while under *nonintegration* agent A chooses q_A and agent B chooses q_B.

Since the types θ_A and θ_B are private information, the outcome at stage 3a cannot depend *directly* on the types. However, the agents' messages may reveal their types. By the revelation principle, we can assume Γ is an incentive compatible revelation mechanism. Thus, at stage 3a, each agent i will simply announce his type θ_i. Incentive compatibility guarantees that truthtelling is an equilibrium. For each revealed type profile $\theta = (\theta_A, \theta_B)$ the mechanism Γ implements an outcome, i.e., a decision profile and a transfer profile. Decisions and transfers are contractible in the sense that if both agent agree to participate in Γ then the outcome at stage 3a is final and not subject to moral hazard or renegotiation. An incentive compatible revelation mechanism which always implements the first best (surplus maximizing) decision profile is said to be a *first best mechanism.*

At stage 2, each agent has an outside option, namely, to refuse to participate in Γ and move to stage 3b. For all types to participate, Γ must satisfy interim individual rationality (IR) constraints: each type must expect to get at least his reservation payoff. The reservation payoff depends on who has the right to make the decisions q_A and q_B at stage 3b. Under integration, agent A chooses both q_A and q_B at stage 3b.[6] Under nonintegration, at stage 3b agent A chooses q_A and agent B chooses q_B simultaneously and independently. (No monetary transfers are ever made at stage 3b.) If a first best mechanism Γ satisfies all interim IR constraints, then Γ is said to be *individually rational.* (By individual rationality, we always mean interim IR.) If an individually rational first-best mechanism Γ exists, then the mediator will propose it, assuming he is instructed to maximize the total surplus. If such Γ exists with integration but not with nonintegration, then integration is chosen at stage 0. Conversely, if such Γ exists with nonintegration but not with integration, then nonintegration is chosen.[7]

3 The First Best

The *social surplus* is the sum of agent A's and agent B's payoff. We represent the social surplus in the following matrix. The row indicates q_A and the

[6]By the symmetry of the agents, it doesn't matter if it is A or B who has the right to make both decisions under integration. For convenience, we assume it is agent A. Agent B does not have the option to "quit the game" at stage 3b, but must remain and accept whatever decision agent A makes. To justify this, assume remaining in the relationship is always better than quitting. Allowing an "exit" option would complicate the exposition without generating any new insights.

[7]If such Γ exists neither for integration nor for nonintegration, then a second-best analysis must be performed to determine which allocation of control rights is preferred. However, as the logic behind the trade-off between integration and nonintegration is revealed by the first best analysis, we omit the much more complex second-best analysis.

column indicates q_B.

$$\begin{array}{ccc} & Up & Down \\ Up & \theta_A + \theta_B + 2\gamma & \theta_A \\ Down & \theta_B & 2\gamma \end{array} \tag{2}$$

This matrix highlights the fact that neither ρ nor (t_A, t_B) matter for social surplus.

For any type profile $\theta = (\theta_A, \theta_B)$, the *first best* (socially optimal) decision profile is denoted $q^*(\theta) = (q_A^*(\theta), q_B^*(\theta))$. By definition, $q^*(\theta)$ maximizes the social surplus, i.e., it selects the biggest number in the matrix (2).[8] Figure 1 illustrates the first best.

Let us consider the first best decision profile from the point of view of agent A, who knows θ_A but not θ_B. (The situation for agent B is symmetric.)

Case 1: $\theta_A \leq -2\gamma$. In this case, the second row of matrix (2) dominates the first row, so $q_A^*(\theta) =$*Down*, regardless of θ_B. The socially optimal q_B does depend on θ_B. Specifically, $q_B^*(\theta) =$*Down* if $\theta_B < 2\gamma$ and $q_B^*(\theta) =$*Up* if $\theta_B > 2\gamma$.

Case 2: $-2\gamma < \theta_A < 2\gamma$. In this case, the surplus maximizing decision is always on the main diagonal of matrix (2). Specifically, $q^*(\theta) =$ (*Down*,*Down*) if $\theta_A + \theta_B < 0$, and $q^*(\theta) =$ (*Up*,*Up*) if $\theta_A + \theta_B > 0$.

Case 3: $\theta_A \geq 2\gamma$. In this case, the first row of matrix (2) dominates the second row, so $q_A^*(\theta) =$*Up*, regardless of θ_B. The socially optimal q_B does depend on θ_B. Specifically, $q_B^*(\theta) =$*Down* if $\theta_B < -2\gamma$ and $q_B^*(\theta) =$*Up* if $\theta_B > -2\gamma$.

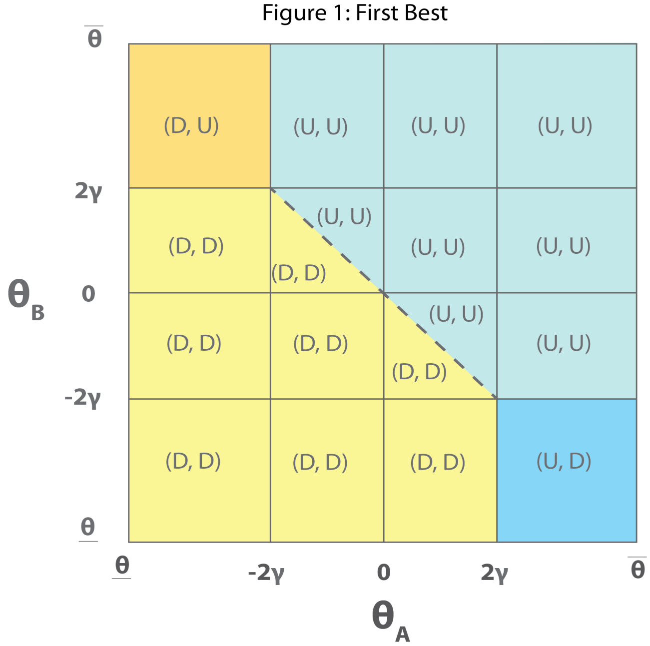
Notice that if $|\theta_A| \geq 2\gamma$ then agent A has such strong idiosyncratic preferences that $q_A^*(\theta)$ is independent of θ_B.

The following matrix indicates the *social surplus generated by agent A*. The row indicates q_A and the column indicates q_B. It differs from matrix (1) because ρ, which is irrelevant for social surplus calculations, has been dropped.

$$\begin{array}{ccc} & Up & Down \\ Up & \theta_A + \gamma & \theta_A \\ Down & 0 & \gamma \end{array}$$

[8] If two numbers are the same, either decision profile can be chosen, but this will happen with probability zero.

Figure 1: First Best



We will calculate the *ex ante* expected social surplus generated by agent A under the first best decision rule, denoted S_A. With probability $F(-2\gamma)$, Case 1 applies to θ_A. In this case, $q^*(\theta) = (Down, Down)$ if $\theta_B < 2\gamma$, and then the social surplus generated by agent A is γ; but $q^*(\theta) = (Down, Up)$ if $\theta_B > 2\gamma$ and then agent A generates no social surplus. Now $\theta_B < 2\gamma$ with probability $F(2\gamma)$, so corresponding to Case 1, S_A contains the term $F(-2\gamma)F(2\gamma)\gamma$. Considering the other two cases in a similar way, we obtain

$$
\begin{aligned}
S_A &= F(-2\gamma)F(2\gamma)\gamma + \int_{-2\gamma}^{2\gamma} (\gamma + \theta_A F(\theta_A))\, dF(\theta_A) + \int_{2\gamma}^{\overline{\theta}} (\theta_A + \gamma F(2\gamma))\, dF(\theta_A) \\
&= 2\gamma \left[(1 - F(2\gamma))F(2\gamma) + F(2\gamma) - \frac{1}{2} \right] + \int_{-2\gamma}^{2\gamma} \theta_A F(\theta_A) dF(\theta_A) \\
&\quad + \int_{2\gamma}^{\overline{\theta}} \theta_A dF(\theta_A)
\end{aligned}
$$

This calculation used Assumption 2. Integration by parts yields

$$\int_{-2\gamma}^{2\gamma} sF(s)dF(s) = \frac{1}{2}\int_{-2\gamma}^{2\gamma} sd(F(s))^2$$

$$= \gamma\left[(F(2\gamma))^2 + (1 - F(2\gamma))^2\right] - \frac{1}{2}\int_{-2\gamma}^{2\gamma}(F(s))^2 ds.$$

Using this, we obtain a simplified expression:

$$S_A = 2F(2\gamma)\gamma - \frac{1}{2}\int_{-2\gamma}^{2\gamma}(F(s))^2 ds + \int_{2\gamma}^{\overline{\theta}} sdF(s). \tag{3}$$

The two agents are perfectly symmetric, so ex ante the expected social surplus under the first best is $S = 2S_A$. Explicitly, we have

$$S = 4\gamma F(2\gamma) - \int_{-2\gamma}^{2\gamma}(F(s))^2 ds + 2\int_{2\gamma}^{\overline{\theta}} sdF(s). \tag{4}$$

4 Individually Rational First Best Mechanisms

4.1 Incentive Compatibility

From now on, we will consider a first best mechanism Γ. The results from Section 3 will be used to compute any type's expected payoff from Γ, assuming all types participate in it (later, we will check if interim IR constraints hold). The agents are symmetric, so it suffices to consider agent A.

Let

$$t(\theta_A) \equiv \int_{\underline{\theta}}^{\overline{\theta}} t_A(\theta_A, \tilde{\theta}_B)dF(\tilde{\theta}_B)$$

denote agent A's expected transfer when his type is θ_A, and let

$$u_A(\theta_A) \equiv \int_{\underline{\theta}}^{\overline{\theta}} v_A(q^*(\theta_A, \tilde{\theta}_B), \theta_A)dF(\tilde{\theta}_B) + t(\theta_A)$$

denote type θ_A's expected payoff. The *ex ante* expected payoff for agent A is

$$V_A \equiv \int_{\underline{\theta}}^{\overline{\theta}} u_A(\theta_A)dF(\theta_A).$$

As explained in Section 3, there are three cases.

Case 1: $\theta_A \leq -2\gamma$. Then $q_A = Down$. Agent A gets γ with probability $F(2\gamma)$ and $-\rho$ with probability $1 - F(2\gamma)$. Thus, agent A's expected payoff is

$$u_A(\theta_A) = t(\theta_A) + \gamma F(2\gamma) - \rho(1 - F(2\gamma)) = t(\theta_A) - \rho + (\gamma + \rho)F(2\gamma). \quad (5)$$

Case 2: $-2\gamma < \theta_A < 2\gamma$. Then the decision is $(Down, Down)$ if $\theta_B < -\theta_A$, and (Up, Up) otherwise, so agent A expects

$$u_A(\theta_A) = t(\theta_A) + \gamma + \theta_A(1 - F(-\theta_A))) = t(\theta_A) + \gamma + \theta_A F(\theta_A). \quad (6)$$

using Assumption 2.

Case 3: $\theta_A \geq 2\gamma$. Then $q_A = Up$. Agent A gets γ with probability $1 - F(-2\gamma) = F(2\gamma)$, and ρ with probability $1 - F(2\gamma)$. Thus, agent A expects

$$u_A(\theta_A) = t(\theta_A) + \theta_A + \gamma F(2\gamma) + \rho(1 - F(2\gamma)) = t(\theta_A) + \rho + \theta_A + (\gamma - \rho)\, F(2\gamma). \quad (7)$$

Incentive compatibility imposes restrictions on the transfer function. In Case 1, the decision profile is independent of θ_A, so the expected transfer must equal some constant t_A for all such types, and then from (5) the expected payoff is constant as well. Thus, for all $\theta_A < -2\gamma$ we have $t(\theta_A) = t_A$ and

$$u_A(\theta_A) = \underline{u}_A \equiv t_A - \rho + (\gamma + \rho)F(2\gamma). \quad (8)$$

In Case 2, the arguments of Myerson and Satterthwaite [17] imply that u_A is differentiable almost everywhere, and $u'_A(\theta_A) = F(\theta_A)$ for θ_A such that $-2\gamma < \theta_A < 2\gamma$. This implies that

$$u_A(\theta_A) = \underline{u}_A + \int_{-2\gamma}^{\theta_A} F(s)ds$$

if $-2\gamma < \theta_A < 2\gamma$.

In Case 3, the arguments of Myerson and Satterthwaite [17] imply $u'_A(\theta_A) = 1$. This implies that

$$\begin{aligned} u_A(\theta_A) &= u_A(2\gamma) + \theta_A - 2\gamma = \underline{u}_A + \int_{-2\gamma}^{2\gamma} F(s)ds + \theta_A - 2\gamma \\ &= \underline{u}_A + \theta_A \end{aligned} \quad (9)$$

if $\theta_A > 2\gamma$. To obtain the final expression in (9), we used the fact that Assumption 2 implies

$$\int_{-2\gamma}^{2\gamma} s dF(s) = 0$$

and

$$\begin{aligned}\int_{-2\gamma}^{2\gamma} F(s)ds &= 2\gamma F(2\gamma) - (-2\gamma)F(-2\gamma) - \int_{-2\gamma}^{2\gamma} s dF(s) \\ &= 2\gamma F(2\gamma) + 2\gamma(1 - F(2\gamma)) - \int_{-2\gamma}^{2\gamma} s dF(s) \\ &= 2\gamma.\end{aligned}$$

Now we combine the three cases to compute the *ex ante* expected payoff for agent A, denoted V_A. Using integration by parts, we get

$$\begin{aligned}V_A &= t_A - \rho + (\gamma+\rho)F(2\gamma) + \int_{-2\gamma}^{2\gamma}\left[\int_{-2\gamma}^{\theta_A} F(s)ds\right] dF(\theta_A) + \int_{2\gamma}^{\overline{\theta}} \theta_A dF_A(\theta_A) \\ &= t_A - \rho + (3\gamma+\rho)F(2\gamma) - \int_{-2\gamma}^{2\gamma} (F(\theta_A))^2 d\theta_A + \int_{2\gamma}^{\overline{\theta}} \theta_A dF_A(\theta_A).\end{aligned}$$

This calculation used the fact that

$$\begin{aligned}\int_{-2\gamma}^{2\gamma}\left[\int_{-2\gamma}^{\theta_A} F(s)ds\right] dF(\theta_A) &= F(2\gamma)\int_{-2\gamma}^{2\gamma} F(s)ds - \int_{-2\gamma}^{2\gamma} (F(\theta_A))^2 d\theta_A \\ &= F(2\gamma)2\gamma - \int_{-2\gamma}^{2\gamma} (F(\theta_A))^2 d\theta_A.\end{aligned}$$

Symmetry implies that agent B's ex ante expected payoff is

$$V_B = t_B - \rho + (3\gamma+\rho)F(2\gamma) - \int_{-2\gamma}^{2\gamma} (F(\theta_B))^2 d\theta_B + \int_{2\gamma}^{\overline{\theta}} \theta_B dF(\theta_B)$$

The sum of the ex ante expected payoffs is

$$\begin{aligned}& V_A + V_B \\ = \; & t_A + t_B - 2\rho + (6\gamma + 2\rho)F(2\gamma) \\ & -2\int_{-2\gamma}^{2\gamma} (F(s))^2 ds + 2\int_{2\gamma}^{\overline{\theta}} s dF(s). \qquad (10)\end{aligned}$$

Budget balance requires that the sum of the expected payoffs equals the expected social surplus:

$$V_A + V_B = S. \tag{11}$$

Above, we showed that all types $\theta_i < -2\gamma$ get the same expected transfer $t(\theta_i) = t_i$, and their expected payoff from participating in Γ is

$$\underline{u}_i \equiv t_i - \rho + (\gamma + \rho)F(2\gamma).$$

Substituting from (10) and (4) in (11), we obtain:

$$\underline{u}_A + \underline{u}_B \equiv t_A + t_B - 2\rho + 2(\gamma + \rho)F(2\gamma) = \int_{-2\gamma}^{2\gamma} (F(s))^2 \, ds. \tag{12}$$

4.2 The AGV Mechanism

The well-known d'Aspremont and Gérard-Varet [3] (AGV) mechanism is a first best mechanism. Given the revealed types (θ_A, θ_B), the first best decision $q^*(\theta_A, \theta_B)$ is chosen, and agent A gets the AGV transfer

$$t_A(\theta_A, \theta_B) \equiv \int_{\underline{\theta}}^{\overline{\theta}} v_B(q^*(\theta_A, \tilde{\theta}_B), \tilde{\theta}_B) dF(\tilde{\theta}_B)$$
$$- \int_{\underline{\theta}}^{\overline{\theta}} v_A(q^*(\tilde{\theta}_A, \theta_B), \tilde{\theta}_A) dF(\tilde{\theta}_A) + k_A \tag{13}$$

where k_A is a constant. Player B's transfer is given by the analogous expression, with a constant k_B. Budget balance requires $k_A + k_B = 0$.

If the agents could engage in comprehensive contracting *ex ante*, before they learn their own types, then they could commit to using an AGV mechanism. Since the AGV mechanism is first best, the constant $k_A = -k_B$ could be chosen to ensure that each agent is willing, ex ante, to participate.[9] With such comprehensive ex ante contracting, the Coase theorem would hold, and the issue of integration versus nonintegration would be moot.[10] In this paper, however, we allow only *interim* contracting. At the interim stage, decisions and transfers are contractible and will be irrevocably specified by the mechanism Γ, provided Γ is accepted by both agents.

[9] Agent i's *ex ante* reservation payoff, denoted r_i, would depend on how decisions are expected to be made if no ex ante contract is signed. Since the AGV mechanism is first best it generates total expected surplus S. By definition of first best, $r_A + r_B \leq S$. Setting k_A appropriately ensures that each agent i gets at least r_i.

[10] The Coase theorem would also hold for contracting under *complete information*, where

4.3 Individual Rationality under Nonintegration

Under nonintegration, if stage 3b is reached then each agent $i \in \{A, B\}$ chooses q_i knowing his own type θ_i but not the other agent's type θ_j. Consider a noncooperative (Bayesian-Nash) equilibrium of this stage 3b game.[11] Since higher types are more inclined to choose *Up*, each agent must use a cutoff strategy. Suppose agent B chooses $q_B =$*Down* if and only if $\theta_B \leq x$, which happens with probability $F(x)$. Then, agent A prefers $q_A =$*Down* if

$$F(x)\gamma - (1 - F(x))\rho \geq \theta_A + F(x)\rho + (1 - F(x))\gamma$$

which is equivalent to

$$\theta_A \leq (2F(x) - 1)\,\gamma - \rho.$$

Thus, if agent B uses cutoff x, agent A's best response is to use the cutoff y defined by

$$y = (2F(x) - 1)\,\gamma - \rho. \tag{14}$$

By Assumption 1, the best response function has slope less than one:

$$\frac{dy}{dx} = 2F'(x)\gamma < 1.$$

Therefore, there is a unique noncooperative equilibrium. By the symmetry of the game, this equilibrium must be symmetric, and it can be found by

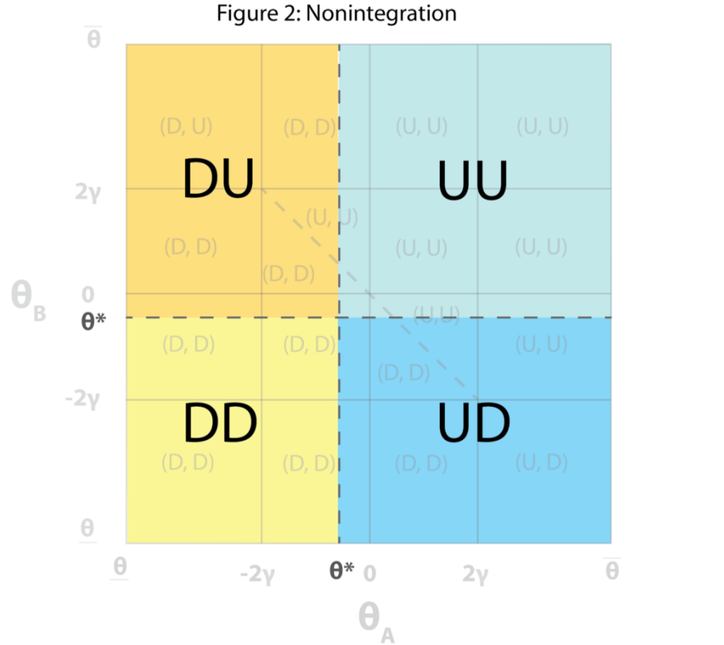
---

each agent knows $\theta = (\theta_A, \theta_B)$. Let agent A make a take-it-or-leave-it offer consisting of a proposed decision profile and a transfer to agent B. If agent B rejects the offer then he gets his reservation payoff $r_B(\theta)$. In subgame perfect equilibrium, agent A proposes $q^*(\theta)$ and a transfer $t_B(\theta)$ such that agent B gets exactly his reservation payoff,

$$v_B(q^*(\theta), \theta_B) + t_B(\theta) = r_B(\theta).$$

[11] To simplify, we make two assumptions about the noncooperative game played at stage 3b: each agent maintains his prior beliefs about the other agent's type, and they do not communicate via cheap-talk. Either assumption could be changed, and the results would be qualitatively the same. Quantitatively the results would change, because the reservation payoffs would change. But the disagreement point would still not be first best, and control rights would still matter for the attainable surplus. The model with cheap talk would be reminiscent of Alonso, Dessein and Matouschek [1] and Rantakari [20], but with a comparison of integration and nonintegration, rather than centralization versus decentralization.

Figure 2: Nonintegration



setting $x = y$ in (14). Thus, in the unique equilibrium, each agent uses the cutoff θ^* defined by

$$\theta^* + \gamma + \rho = 2F(\theta^*)\gamma. \tag{15}$$

It can be checked that $-\gamma - \rho < \theta^* < -\rho$. Figure 2 illustrates the noncooperative equilibrium under nonintegration.

Agent A's noncooperative equilibrium payoff is

$$\gamma F(\theta^*) - \rho(1 - F(\theta^*)) = -\rho + (\gamma + \rho)F(\theta^*) \tag{16}$$

if $\theta_A < \theta^*$ and

$$\theta_A + \gamma + (\rho - \gamma)F(\theta^*) \tag{17}$$

if $\theta_A > \theta^*$.

Since the noncooperative equilibrium is played at stage 3b if Γ is rejected at stage 2, agent A's reservation payoff at stage 2 is given by (16) if $\theta_A < \theta^*$ and (17) if $\theta_A > \theta^*$.

Let Γ be a first best mechanism. Which type of agent A is most reluctant to accept Γ at stage 2? If stage 3b is reached, there is no check on opportunism, and type $\theta_A > \theta^*$ benefits by playing *Up*, thus gaining ρ when $\theta_B < \theta^*$. Since ρ does not count as a gain for social surplus computations, this opportunistic benefit is reduced by Γ. Indeed, consider type θ_A such that $\theta^* < \theta_A \leq 2\gamma$. If he accepts Γ, the first best decision rule will, with some probability, make him forego the opportunistic gain ρ by implementing $q_A =$*Down*. As the benefit from $q_A =$*Down* is decreasing in the type, in the range $[\theta^*, 2\gamma]$ the disadvantage of the first best is the greatest for type 2γ. On the other hand, if $\theta_A > 2\gamma$ then $q_A =$*Up* both in the noncooperative equilibrium and in the first best, so the disadvantage is not any greater for types above 2γ than it is for type 2γ. This intuitive argument suggests that the most difficult interim IR constraint to satisfy is for type 2γ. The following lemma verifies that the argument is correct.

Lemma 1 *Let Γ be a first best mechanism. With nonintegration, agent i's interim IR constraints are satisfied for all his types if the interim IR constraint is satisfied for type $\theta_i = 2\gamma$, which is true if and only if*

$$\underline{u}_i \geq \gamma + (\rho - \gamma)F(\theta^*). \tag{18}$$

The proof of this and subsequent lemmas can be found in the appendix.

4.4 Individual Rationality under Integration

Under integration, if stage 3b is reached, agent A unilaterally chooses the decision profile (q_A, q_B) which is best for himself, given θ_A. That is, he picks the greatest number in the payoff matrix (1), for $i = A$. Notice that agent A's decision doesn't depend on his beliefs about θ_B, since θ_B doesn't influence his payoff.

Since agent A has right to choose the decision profile unilaterally if stage 3b is reached, his reservation payoff at stage 2 is high. If agent B's reservation payoff were correspondingly low, agent B would be willing to pay a high transfer to agent A at stage 3a in order to implement the first best. In this case, an AGV mechanism with a high k_A would be individually rational. But agent B's reservation payoff depends on agent A's incentive to cooperate or rent-seek. In the cooperative case where $\gamma > \rho$, at stage 3b agent A will never behave opportunistically (in the sense of transferring ρ from agent B to himself). In the *rent-seeking case* where $\rho > \gamma$, at stage 3b agent A will sometimes choose $(q_A, q_B) = $ (*Up*,*Down*), a rent-seeking policy which transfers ρ from agent B to himself. Since the IR constraints will differ in the two cases, we treat them separately. We emphasize the key fact that agent B's reservation payoff at stage 2 is quite different in the two cases: high in the cooperative case (making it hard to satisfy the IR constraints) but low in the rent-seeking case (making it easy to satisfy the IR constraints).

4.4.1 The Cooperative Case

Suppose $\rho < \gamma$ and Γ is rejected at stage 2. At stage 3b agent A chooses ($Down$,$Down$) if $\theta_A < 0$ and (Up, Up) otherwise. Therefore, if agent A rejects Γ at stage 2 then his payoff at stage 3b will be

$$\gamma + \max\{0, \theta_A\}. \tag{19}$$

This is agent A's reservation payoff at stage 2. Of course, since agent A does not take θ_B into account, his decision is unlikely to be first best. At stage 3b agent B gets γ if $\theta_A < 0$ and $\theta_B + \gamma$ otherwise, so if agent B rejects Γ at stage 2 his expected payoff is

$$\gamma + (1 - F(0))\,\theta_B = \gamma + \theta_B/2. \tag{20}$$

This is agent B's reservation payoff at stage 2. Figure 3 illustrates what happens at stage 3b.

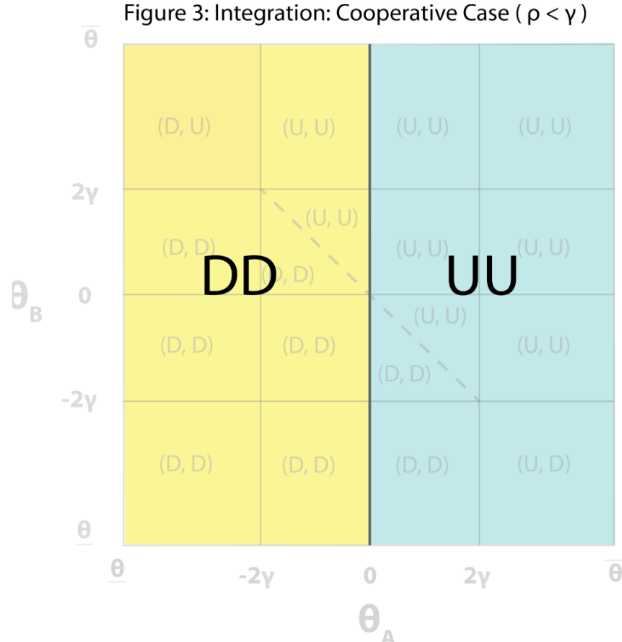
In the cooperative case, agent A will never hold up agent B and take ρ away at stage 3b. In fact, although agent A's choice is not *always* first best, it *is* first best when $\theta_B = 0$, since then only θ_A matters for social efficiency. Type $\theta_B = 0$ knows that the decision profile will be the same whether they reach stage 3a or 3b, so he is not willing to pay anything to participate in Γ.[12] Clearly, the IR constraint for type $\theta_B = 0$ will be difficult to satisfy, and this will be used in Section 6 to show that Γ cannot be individually rational. Here, we prove a preliminary result. Recall that in the first best mechanism Γ the expected payoff for any type $\theta_i < -2\gamma$ is given by

$$\underline{u}_i \equiv t_i - \rho + (\gamma + \rho)F(2\gamma).$$

Lemma 2 *Let Γ be a first best mechanism and suppose $\rho < \gamma$. With integration the following is true. (a) Agent A's interim IR constraints are satisfied*

[12] Type $\theta_B \neq 0$ does care about whether they coordinate on ($Down$,$Down$) or (Up, Up). So the most difficult IR constraint to satisfy is for $\theta_B = 0$.

Figure 3: Integration: Cooperative Case (ρ < γ)



for all his types if the interim IR constraint is satisfied for type $\theta_A = 2\gamma$, *which is true if and only if*

$$\underline{u}_A \geq \gamma. \tag{21}$$

(b) Agent B's interim IR constraints are satisfied for all his types if the interim IR constraint is satisfied for type $\theta_B = 0$, *which is true if and only if*

$$\underline{u}_B \geq \gamma - \int_{-2\gamma}^{0} F(s)ds. \tag{22}$$

4.4.2 The Rent-Seeking Case

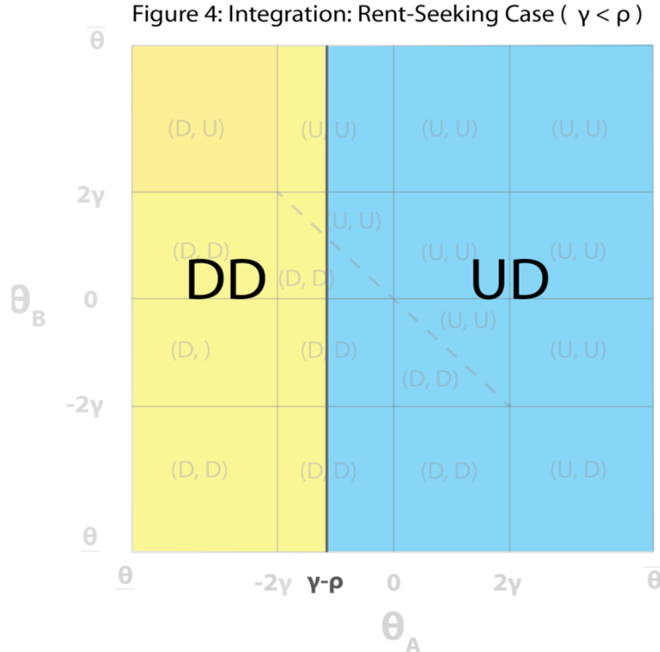
Suppose $\rho > \gamma$ and Γ is rejected at stage 2. At stage 3b, agent A chooses chooses $(Down,Down)$ if $\theta_A < \gamma - \rho$ and $(Up,Down)$ otherwise. Thus, in this case agent A's reservation payoff at stage 2 is

$$\max\{\gamma, \theta_A + \rho\}.$$

At stage 3b agent B receives γ if $\theta_A < \gamma - \rho$ and $-\rho$ otherwise, so agent B's reservation payoff at stage 2 is

$$F(\gamma - \rho)\gamma - \rho(1 - F(\gamma - \rho)) = -\rho + F(\gamma - \rho)(\gamma + \rho).$$

Figure 4: Integration: Rent-Seeking Case (γ < ρ)



Notice that $\gamma - \rho < 0$ so $F(\gamma - \rho) < 1/2$. Figure 4 illustrates the stage 3b rent-seeking policy.

At stage 3b, agent A will always choose $q_B =$*Down* in the rent-seeking case. Since the high types of agent B would derive a large benefit from $q_B =$*Up*, they are eager to accept Γ. In contrast, the low types of B do not mind if $q_B =$*Down*. Indeed, if $\theta_B \leq -2\gamma$ then $q_B =$*Down* in the first best for sure, i.e., q_B will be the same at stages 3a and 3b when $\theta_B \leq -2\gamma$. These types all have the same IR constraint, and it is intuitively clear that this is most difficult IR constraint among agent B's types.

Lemma 3 *Let Γ be a first best mechanism and suppose $\rho > \gamma$. With integration the following is true. (a) Agent A's interim IR constraints are satisfied for all his types if the interim IR constraint is satisfied for type $\theta_A = 2\gamma$, which is true if and only if*

$$\underline{u}_A \geq \rho. \tag{23}$$

(b) Agent B's interim IR constraints are satisfied for all his types if the interim IR constraint is satisfied for type $\theta_B = -2\gamma$, which is true if and only if

$$\underline{u}_B \geq -\rho + F(\gamma - \rho)(\gamma + \rho). \tag{24}$$

5 Implementing the First Best with Nonintegration

Lemma 1 and Equation (12) have the following implication:

Theorem 4 *With nonintegration, there exists an individually rational first best mechanism if and only if*

$$\int_{-2\gamma}^{2\gamma} (F(s))^2 \, ds \geq 2\rho F(\theta^*) + 2\gamma(1 - F(\theta^*)). \tag{25}$$

Proof. The necessity of (25) follows from (12) and Lemma 1.

To prove sufficiency of (25), consider the AGV mechanism with transfers given by (13). Since the mechanism is incentive compatible and the outcome is first best, the results of Section 4 apply. If $k_A = k_B = 0$ then the mechanism is *symmetric*, so $\underline{u}_A = \underline{u}_B$. From (12) we obtain

$$\underline{u}_A = \frac{1}{2} \int_{-2\gamma}^{2\gamma} (F(s))^2 \, ds. \tag{26}$$

Inequality (25) says that (26) exceeds $\gamma + (\rho - \gamma)F(\theta^*)$. Therefore, by Lemma 18, all IR constraints are satisfied. ■

Theorem 4 is hard to interpret because θ^* is an endogenous variable which depends on γ and ρ. The following result helps us express Theorem 4 in terms of γ and ρ.

Lemma 5 *For any γ, there is $\rho^*(\gamma) \in (0, \gamma)$ such that (25) holds if and only if $\rho \leq \rho^*(\gamma)$.*

Combining Theorem 4 and Lemma 5, we get the following result.

Theorem 6 *With nonintegration, there exists an individually rational first best mechanism if and only if $\rho \leq \rho^*(\gamma)$, where $0 < \rho^*(\gamma) < \gamma$.*

To understand this result intuitively, note that Lemma 1 showed that the most difficult IR constraint to satisfy is for type 2γ. The noncooperative nonintegrated equilibrium favors type 2γ inasmuch as he can benefit from the opportunistic benefit ρ. Still, when ρ is low, type 2γ does not need a big transfer to accept Γ. Then, other types would not gain much by pretending to be type 2γ, i.e., the benefits of "haggling" are low. Transfers to all types can be correspondingly low without violating incentive compatibility. Then there is enough surplus at the first best to satisfy all incentive and participation constraints without violating budget balance. But when ρ is high, it takes a big expected transfer to make type 2γ accept Γ. The incentive constraints imply the expected transfer must be high to all types. In terms of the AGV mechanism, we need $k_A > 0$ and $k_B > 0$ which violates budget balance, so the first best cannot be implemented.

We end our analysis of nonintegration by considering the role of Assumption 1. If this assumption does not hold, there can be multiple noncooperative equilibria. Then, there may be some range of ρ such that the IR constraints can be satisfied or not, depending on *which* noncooperative equilibrium is played at stage 3. However, one would expect that if ρ is sufficiently small then *all* equilibria should allow the IR constraint to be satisfied. Conversely, if ρ is sufficiently large then *no* equilibrium should allow the IR constraint to be satisfied.

We can verify this intuition most simply under the assumption that the agents play a symmetric equilibrium because they always exist. Multiple symmetric equilibria correspond to multiple solutions to (15), but all solutions lie between $-\rho - \gamma$ and $-\rho$. If $\rho > \gamma$, then the same arguments as above show that the first best is not implementable under nonintegration, even when there are multiple equilibria. Suppose $\rho < \gamma$ and consider any equilibrium with a cutoff θ^*. The right hand side of (25) tends to 2γ as ρ gets close to γ. Hence, for ρ close to γ, the first best is not implementable under nonintegration. For ρ close to zero, the right hand side of (25) is approximately $2\gamma(1 - F(\theta^*))$. We know $\theta^* > -\rho - \gamma \approx -\gamma$ so

$$2\gamma(1 - F(\theta^*)) < 2\gamma(1 - F(-\gamma)) = 2\gamma F(\gamma)$$

by symmetry. If there is enough probability weight on high types, so

$$\int_{-2\gamma}^{2\gamma} (F(s))^2\, ds > 2\gamma F(\gamma), \tag{27}$$

then (25) holds for sufficiently small ρ. Therefore, the first best is implementable under nonintegration even when there are multiple equilibria, when ρ is small enough. Thus, qualitatively, the results of this section obtain even if Assumption 1 is violated and multiple non-cooperative equilibria exist: for some $\underline{\rho}$, $\overline{\rho} \in (0, \gamma)$ such that $\underline{\rho} < \overline{\rho}$, the first best is implemented by nonintegration if $\rho < \underline{\rho}$, but not if $\rho > \overline{\rho}$.

6 Integration: The Cooperative Case

Suppose $\rho < \gamma$. Then, from Lemma 2 and (12), if the first best is implementable under integration then

$$\underline{u}_A + \underline{u}_B = \int_{-2\gamma}^{2\gamma} (F(s))^2 \, ds \geq 2\gamma - \int_{-2\gamma}^{0} F(s) ds. \tag{28}$$

But this never holds, because the middle term and the third term are equal at $\gamma = 0$ and the derivative of the middle term with respect to γ is always strictly less than the derivative of the third term. Therefore, the inequality in (28) is violated for all $\gamma > 0$. Thus, we have the following result:

Theorem 7 *If $\rho < \gamma$, then with integration no first best mechanism is individually rational.*

Thus, it is impossible to implement the first best with integration in the cooperative case. To see this intuitively, recall that since agent A has control rights if stage 3b is reached, his reservation payoff at stage 2 is high. If agent B's reservation payoff were correspondingly low, then an AGV mechanism with a large positive k_A and correspondingly large negative k_B would be individually rational. However, as argued above, type $\theta_B = 0$ knows that the decision profile will be first best whether they reach stage 3a or stage 3b, so he will certainly reject an AGV mechanism with a large negative k_B. But agent A will insist that k_A should be large and positive on account of his advantageous bargaining position (he has the right to choose both q_A and q_B). Since budget balance forces $k_B = -k_A$ the first best cannot be implemented.

Notice that if $\theta_B \neq 0$ then at stage 3b agent B may suffer relative to the first best outcome, as agent A's unilateral decision may cause agent B

to lose surplus if $\theta_B < 0$ or not capture enough if $\theta_B > 0$. So if agent B has strong idiosyncratic preferences, he would in principle be willing to pay a large transfer to reach the first best. But, because his type is private information, agent B can haggle, i.e., pretend to be type $\theta_B = 0$ in order to avoid the transfer. Types $\theta_B \neq 0$ must receive information rents to preserve incentive compatibility. These rents limit the ability to extract transfers from agent B's types and prevent implementation of the first best.

Combining Theorems 7 and 6 we get:

Theorem 8 Cooperative Case. *Suppose $\gamma > \rho$. Then, with integration no first best individually rational mechanism exists; with nonintegration such a mechanism exists if and only if $\rho \leq \rho^*(\gamma)$, where $0 < \rho^*(\gamma) < \gamma$.*

7 Integration: The Rent-Seeking Case

Suppose $\rho > \gamma$. From Theorem 6, as $\rho^*(\gamma) < \gamma$, we already know that the first best cannot be implemented under nonintegration. High rents imply that the benefits of haggling are high under nonintegration and this creates inefficiency. In contrast, there are circumstances where the first best can be implemented under integration:

Theorem 9 *If $\rho > \gamma$, then with integration a first best individually rational mechanism exists if and only if*

$$\int_{-2\gamma}^{2\gamma} (F(s))^2 \, ds \geq F(\gamma - \rho)(\gamma + \rho). \tag{29}$$

Proof. The necessity of (29) follows from Lemma 3 and (12). To prove sufficiency of (29), consider the AGV mechanism with transfers given by (13). Since the mechanism is incentive compatible and the outcome is first best, the results of Section 4 apply. Choose k_A such that $\underline{u}_A = \rho$. By Lemma 3, all of agent A's IR constraints are satisfied. From (12) and (29),

$$\underline{u}_B = \int_{-2\gamma}^{2\gamma} (F(s))^2 \, ds - \rho \geq F(\gamma - \rho)(\gamma + \rho) - \rho.$$

Again by Lemma 3, all of agent B's IR constraints are satisfied as well. ■

To interpret Theorem 9, notice that (29) holds in the following three scenarios: (i) when the idiosyncratic preference is relatively unimportant; (ii) when ρ is only slightly bigger than γ; (iii) when ρ is much bigger than γ.

To prove (i), define $z \equiv \rho - \gamma > 0$. We make θ_i less important, without changing the relative size of ρ and γ, by simultaneously increasing ρ and γ, while keeping $z \equiv \rho - \gamma$ fixed. Now (29) can be written as

$$\int_{-2\gamma}^{2\gamma} (F(s))^2 \, ds \geq F(-z)(2\gamma + z).$$

Raising γ increases the left hand side at the rate

$$2\left(F(2\gamma)\right)^2 + 2\left(F(-2\gamma)\right)^2 = 2\left(F(2\gamma)\right)^2 + 2\left(1 - F(2\gamma)\right)^2$$

which always exceeds 1 and, in fact, will be close to 2 for γ large. But the right hand side goes up only at the rate $2F(-z) < 1$ because $-z < 0$ and $F(0) = 1/2$. Therefore, (29) is bound to hold for γ large enough, for any given z.

To prove (ii), note that (29) holds with strict inequality when $\rho = \gamma$, because the left side lies strictly between γ and 2γ while the right side equals γ when $\rho = \gamma$. Thus (29) also holds for ρ slightly above γ.

To prove (iii), note that $F(\gamma - \rho) = 0$ whenever $\gamma - \rho < \underline{\theta}$. Therefore, (29) holds if $\rho > \gamma - \underline{\theta}$.

To visualize inequality (29), suppose F is the normal distribution with mean 0 and variance 1. In Figure 5, the light blue shaded (or lightly shaded) area is the set of (γ, ρ) such that $\rho > \gamma$ and the inequality (29) holds. Thus, if (γ, ρ) belongs to the light blue shaded area, the first best can be implemented under integration. The dark (blue) curve is the set of (γ, ρ) such that there is equality in (29). If both γ and ρ increase simultaneously then we move into the light blue area (property (i)); the light blue area includes all (γ, ρ) slightly above the 45 degree line (property (ii)); and if ρ is increased while γ is fixed, we move into the light blue area (property (iii)).

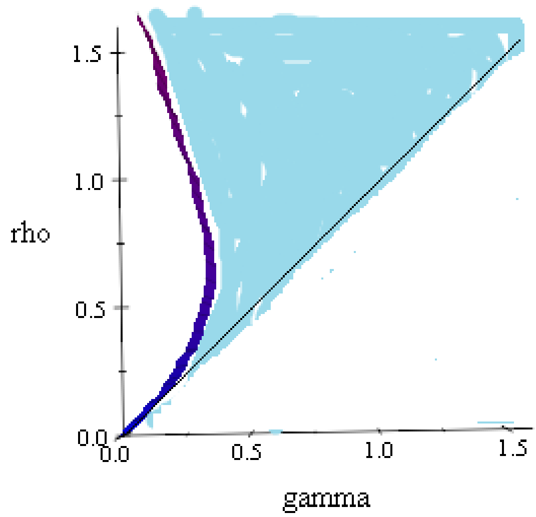


Figure 5

To understand why is it possible to implement the first best with integration in the rent-seeking case, notice that if agent B rejects Γ at stage 2, there is a chance that he will be held up at stage 3b (the outcome will be (Up,$Down$) and agent B loses ρ). This imposes an expected cost on all of player B's types; in terms of the AGV mechanism, he is willing to agree to set $k_A = -k_B > 0$. Haggling (misrepresenting information) is not a serious problem: because all of agent B's types will suffer from hold up if negotiations break down, there is no type which has to be given a large expected payoff to participate, hence agent B does not have much to gained from misrepresenting his type. Surplus is destroyed if stage 3b is reached because actions are not efficiently coordinated, and the extra surplus generated by the first best can be sufficient to satisfy both agents' individual rationality constraints.

Combining Theorem 6 and Theorem 9, we have the following result:

Theorem 10 Rent-Seeking Case. *Suppose $\gamma < \rho$. Then, with nonintegration no first best individually rational mechanism exists; with integration such* *a mechanism exists if and only if*

$$\int_{-2\gamma}^{2\gamma} (F(s))^2 \, ds \geq F(\gamma - \rho)(\gamma + \rho).$$

8 Summary

Grossman and Hart [6] showed how the allocation of control rights can impact noncontractible ex ante investments. In our model, there are no ex ante investments, but control rights matter because of information asymmetries. In reality, information asymmetries seem to be pervasive. Pisano [19] observes that in the biotechnology industry:

> The firm in charge of R&D will accumulate asymmetric information on the technology; likewise, the partner in charge of marketing will gain asymmetric information on the technology's commercial potential. Strategic misrepresentation of new information by either party is a possibility (Pisano [19]).

In the biotech industry, various ownership structures exist: one firm may own a controlling share in a separate jointly controlled venture or become a large but minority shareholder in its trading partner; or two firms may have fifty-fifty ownership and control specific parts of the joint venture, say with one controlling decisions related to R&D and the other decisions related to marketing (Pisano [19]).

Our results show how the maximum surplus that can be achieved by interim negotiations depends on the allocation of control rights. With nonintegration, each agent has a moderate degree of bargaining power when decisions are taken, because each agent has some decision rights if negotiations fail. But there is scope for "haggling" in the sense of preference falsification. Each agent has some type who expects to benefit from rent-seeking, and therefore have high reservation payoff. But the strong bargaining position of this type spills over to other types in the process of haggling, where they may *pretend* to be this type. Therefore, if ρ is large (rent-seeking is very profitable) then the first best cannot be implemented with nonintegration: to both induce the type who benefits the most from rent-seeking to participate in the first best *and* prevent other types from haggling would require more surplus than is available. This is a formalization of Williamson's [23] argument that haggling by nonintegrated agents may destroy surplus.

Williamson [23] argued that integration reduces the cost of haggling. In our model, either integration or nonintegration may have lower cost, depending on the parameters. Under integration, agent A chooses both q_A and q_B if negotiations break down, so agent A has a high reservation payoff. If agent B's reservation payoff is correspondingly low, then implementation of the first best will be possible, for agent B will be willing to make a monetary transfer to agent A in order to choose the first best decision. However, agent B's reservation payoff depends on agent A's relative incentives to cooperate or rent-seek. In the rent-seeking case, agent B's reservation payoff is low because he expects to be held up. Therefore, if ρ is large (compared to γ), the first best can be implemented with integration. Intuitively, the disutility the boss imposes on the subordinate if negotiations break down is sufficient to outweigh the benefit to the boss from doing so, and this makes it possible to agree on the first best. In the cooperative case, however, we showed that agent B's type 0 has a high reservation payoff. The possibility of haggling means agent B's other types can pretend to be type 0, giving them significant bargaining power as well. Thus, the mechanism Γ cannot specify that agent B pays a large transfer to agent A. But without such a transfer, agent A prefers to exercise his control rights unilaterally. Therefore, if ρ is small (compared to γ), the first best cannot be implemented with integration.

For the case where F is the normal distribution with mean 0 and variance 1, Figure 6 summarizes the results for both integration and nonintegration.

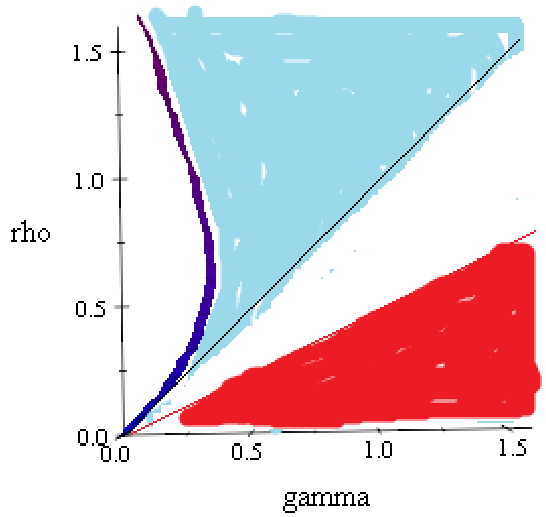


Figure 6

The light blue or lightly shaded area is the set of (γ, ρ) such that with integration a first best individually rational mechanism exists. The red or darkly shaded area is the set of (γ, ρ) such that with nonintegration a first best individually rational mechanism exists (that is, in this area we have $\rho < \rho^*(\gamma)$). In the uncolored area, neither integration nor nonintegration can implement the first best. But in this area either integration or nonintegration would yield the highest surplus (albeit less than first best); so if we considered the second best, the colored areas would extend to cover the whole figure. In the light blue area integration would dominate nonintegration, while in the red area nonintegration would dominate integration. The precise boundary between the two areas, which of course would lie in the uncolored part of Figure 6, could be calculated – but this calculation would not add any new insights.

References

[1] Ricardo Alonso, Wouter Dessein and Niko Matouschek (2008): "When Does Coordination Require Centralization?," *American Economic Re-*

view 98 (1), 145-179.

[2] George Baker, Robert Gibbons and Kevin Murphy (1999) “Informal Authority in Organizations” *Journal of Law, Economics and Organization* 15 (1): 56-73.

[3] Claude d’Aspremont and L. A. Gérard-Varet (1979): “Incentives and Incomplete Information,” *Journal of Public Economics*, 11, 25-45.

[4] Peter Cramton, Robert Gibbons, and Paul Klemperer (1987): “Dissolving a Partnership Efficiently,” *Econometrica*, 55, 615–632.

[5] Robert Gibbons (2005): “Four Formal(izable) Theories of the Firm,” *Journal of Economic Behavior and Organization*, 58, 202-47.

[6] Sanford J. Grossman and Oliver D. Hart (1986): “The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration,” *Journal of Political Economy*, 94, 691–719.

[7] Maija Halonen (2002): “Reputation and The Allocation of Authority,” *Economic Journal*, 112, 539–558.

[8] Oliver Hart and Bengt Holmström (2010): “A Theory of Firm Scope”, *Quarterly Journal of Economics*, May, 125: 483-513.

[9] Benjamin Klein, Robert Crawford, Armen Alchian (1978): “Vertical Integration, Appropriable Rents and the Competitive Contracting Process,” *Journal of Law and Economics*, 21, 297-326.

[10] Oliver D. Hart and John Moore (1990), “Property Rights and the Nature of the Firm,” *Journal of Political Economy*, 98, 1119–1158.

[11] Vijay Krishna and Motty Perry (2000): “Efficient Mechanism Design,” mimeo, Penn State.

[12] Jean-Jacques Laffont and Eric Maskin (1979): “A differentiable approach to expected utility-maximizing mechanisms”, in J.-J. Laffont (ed.), *Aggregation and Revelation of Preferences*. North-Holland, Amsterdam.

[13] Louis Makowski and Claudio Mezzetti (1994), "Bayesian and Weakly Robust First Best Mechanisms: Characterizations," *Journal of Economic Theory*, 64, 500–519.

[14] Eric Maskin and Jean Tirole (1999), "Unforeseen Contingencies and Incomplete Contracts," *Review of Economic Studies*, 66:83-114.

[15] Niko Matouschek (2004): "Ex Post Inefficiencies in a Property Rights Theory of the Firm," *Journal of Law, Economics, & Organization*, 20 (1), 125-147.

[16] Paul Milgrom and John Roberts (1990): "Bargaining costs, influence costs, and the organization of economic activity," in J. E. Alt and K. A. Shepsle (eds.) *Perspectives on positive political economy*, Cambridge University press, 57-89.

[17] Roger B. Myerson and Mark A. Satterthwaite (1983), "Efficient Mechanisms for Bilateral Trading," *Journal of Economic Theory*, 29, 265–281.

[18] Zvika Neeman (1999): "Property Rights and Efficiency of Voluntary Bargaining Under Asymmetric Information," *Review of Economic Studies*, 66, 679–691.

[19] Gary Pisano (1989): "Using Equity Participation to Support Exchange: Evidence from the Biotechnology Industry," *Journal of Law, Economics and Organization*, 5, 109-126.

[20] Heikki Rantakari (2008): "Governing Adaptation," *Review of Economic Studies*, 75(4), 1257-1285.

[21] Steven R. Williams (1999): "A Characterization of Efficient, Bayesian Incentive Compatible Mechanisms," *Economic Theory*, 14, 155–180.

[22] Oliver E. Williamson (1971): "The Vertical Integration of Production: Market Failure Considerations," *American Economic Review*, 61, 112-123.

[23] Oliver E. Williamson (1975): *Markets and Hierarchies: Analysis and Antitrust Implications*, Free Press, New York, N.Y.

[24] Oliver E. Williamson (1979) "Transaction-Cost Economics: The Governance of Contractual Relations" *Journal of Law and Economics* Vol. 22, No. 2, 233-261

[25] Oliver E. Williamson (1983): "Credible Commitments: Using Hostages To Support Exchange," *American Economic Review*, 73(4), 519-540.

[26] Oliver E. Williamson (1985): *The Economic Institutions of Capitalism*, Free Press, New York, N.Y.

9 Appendix

Proof of Lemma 1. By symmetry, it suffices to consider agent A. The IR constraints require that agent A's payoff under Γ, which is either (5), (6) or (7) depending on θ_A, exceeds the reservation payoff, which is either (16) or (17), again depending on θ_A. Thus, we consider the possible cases that can occur.

If $\theta_A \leq \min\{\theta^*, -2\gamma\}$, then the IR constraint is that (5) should be no less than (16), that is,

$$u_A(\theta_A) = \underline{u}_A \geq -\rho + (\gamma + \rho)F(\theta^*).$$

If $\min\{\theta^*, -2\gamma\} \leq \theta_A \leq \max\{\theta^*, -2\gamma\}$, then the IR constraint is

$$u_A(\theta_A) = \underline{u}_A + \int_{-2\gamma}^{\theta_A} F(s)ds \geq -\rho + (\gamma + \rho)F(\theta^*).$$

If $\max\{\theta^*, -2\gamma\} \leq \theta_i \leq 2\gamma$, then the IR constraint is

$$u_A(\theta_A) = \underline{u}_A + \int_{-2\gamma}^{\theta_A} F(s)ds \geq \theta_A + \gamma + (\rho - \gamma)F(\theta^*).$$

If $\theta_i \geq 2\gamma$, then the IR constraint is

$$u_A(\theta_A) = \underline{u}_A + \theta_A \geq \theta_A + \gamma + (\rho - \gamma)F(\theta^*). \quad (30)$$

Notice that

$$\gamma + (\rho - \gamma)F(\theta^*) > -\rho + (\gamma + \rho)F(\theta^*)$$

using (15) and the fact that $\theta^* < 0$. Therefore, the individual rationality constraint for $\theta_A = 2\gamma$, which can be written as

$$\underline{u}_A \geq \gamma + (\rho - \gamma)F(\theta^*),$$

implies all the others. ■

Proof of Lemma 2. The proof is analogous to the proof of Lemma 1. The IR constraints require that the payoff under the mechanism, which is either (5), (6) or (7) depending on the agent's type, exceeds the reservation payoff, which is (19) for agent A and (20) for agent B. Thus, we consider the possible cases that can occur. First, we consider agent A.

If $\theta_A \leq -2\gamma$ then the IR constraint is that (5) should be no less than γ, that is,

$$u_A(\theta_A) = \underline{u}_A \geq \gamma.$$

If $-2\gamma \leq \theta_A \leq 0$, then the IR constraint is

$$u_A(\theta_A) = \underline{u}_A + \int_{-2\gamma}^{\theta_A} F(s)ds \geq \gamma.$$

If $0 \leq \theta_A < 2\gamma$, then the IR constraint is

$$u_A(\theta_A) = \underline{u}_A + \int_{-2\gamma}^{\theta_A} F(s)ds \geq \gamma + \theta_A.$$

If $\theta_A \geq 2\gamma$, then the IR constraint is

$$u_A(\theta_A) = \underline{u}_A + \theta_A \geq \gamma + \theta_A.$$

It is easy to see that all of agent A's IR constraints are satisfied if the IR constraint holds for type $\theta_A = 2\gamma$, which is (21).

For agent B, the individual rationality constraints are as follows.

If $\theta_B \leq -2\gamma$, then the IR constraint is

$$u_B(\theta_B) = \underline{u}_B \geq \gamma + \frac{1}{2}\theta_B.$$

If $-2\gamma \leq \theta_B \leq 2\gamma$, then the IR constraint is

$$u_B(\theta_B) = \underline{u}_B + \int_{-2\gamma}^{\theta_B} F(s)ds \geq \gamma + \frac{1}{2}\theta_B.$$

If $\theta_B \geq 2\gamma$, then the IR constraint is

$$u_B(\theta_B) = \underline{u}_B + \theta_B \geq \gamma + \frac{1}{2}\theta_B.$$

It is easy to check that all of agent B's IR constraints are satisfied if the IR constraint holds for type $\theta_B = 0$, which is (22). ■

Proof of Lemma 3. The proof is analogous to the proofs of Lemmas 1 and 2. Suppose first that $\gamma - \rho < -2\gamma$. Then, the individual rationality constraints for agent A are as follows.

If $\theta_A \leq \gamma - \rho$, then the IR constraint is

$$u_A(\theta_A) = \underline{u}_A \geq \gamma. \tag{31}$$

If $\gamma - \rho \leq \theta_A \leq -2\gamma$, then the IR constraint is

$$u_A(\theta_A) = \underline{u}_A \geq \theta_A + \rho. \tag{32}$$

If $-2\gamma \leq \theta_A \leq 2\gamma$, then the IR constraint is

$$u_A(\theta_A) = \underline{u}_A + \int_{-2\gamma}^{\theta_A} F(s)ds \geq \theta_A + \rho. \tag{33}$$

If $\theta_A \geq 2\gamma$, then the IR constraint is

$$u_A(\theta_A) = \underline{u}_A + \theta_A \geq \theta_A + \rho.$$

If instead $\gamma - \rho > -2\gamma$ then (31) applies when $\theta_A \leq -2\gamma$, (33) applies when $\gamma - \rho \leq \theta_A \leq 2\gamma$, and (32) is replaced by the following IR constraint for the case $-2\gamma \leq \theta_A \leq \gamma - \rho$:

$$u_A(\theta_A) = \underline{u}_A + \int_{-2\gamma}^{\theta_A} F(s)ds \geq \gamma.$$

In either case, it can be checked that all of agent A's individual rationality constraints hold if and only if the individual rationality constraint holds for type $\theta_A = 2\gamma$, which is (23).

For agent B, the individual rationality constraints are as follows.

If $\theta_B \leq -2\gamma$ then the IR constraint is

$$\underline{u}_B \geq -\rho + F(\gamma - \rho)(\gamma + \rho).$$

If $-2\gamma \leq \theta_B \leq 2\gamma$, then the IR constraint is

$$u_B(\theta_B) = \underline{u}_B + \int_{-2\gamma}^{\theta_B} F(s)ds \geq -\rho + F(\gamma - \rho)(\gamma + \rho).$$

If $\theta_B \geq 2\gamma$, then the IR constraint is

$$u_B(\theta_B) = \underline{u}_B + \theta_B \geq -\rho + F(\gamma - \rho)(\gamma + \rho).$$

It can be checked that all of agent B's individual rationality constraints hold if and only if the individual rationality constraint holds for $\theta_B = -2\gamma$, which is (24). ■

Proof of Lemma 5. The left side of (25) can be written as

$$\begin{aligned}
\int_{-2\gamma}^{2\gamma} (F(s))^2 \, ds &= \int_{-2\gamma}^{0} (1 - F(-s))^2 \, ds + \int_{0}^{2\gamma} (F(s))^2 \, ds \\
&= \int_{0}^{2\gamma} \left[(F(s))^2 + (1 - F(s))^2 \right] ds.
\end{aligned}$$

Since $1/2 < (F(s))^2 + (1 - F(s))^2 < 1$, this expression lies strictly between γ and 2γ. Therefore, if $\rho \geq \gamma$, it is impossible to satisfy (25). If ρ is close to 0, then θ^* is close to 0 and the right hand side of (25) is close to γ. If ρ is close to γ, then the right hand side of (25) is close to 2γ.

We now claim that the right hand side of (25) strictly increases from γ to 2γ as ρ increases from 0 to γ. As the left hand side lies strictly between γ and 2γ and is independent of ρ, this claim will prove the lemma.

From (15), we get

$$\frac{d\theta^*}{d\rho} = \frac{1}{2F'(\theta^*)\gamma - 1} < 0.$$

The derivative of the right side of (25) with respect to ρ is

$$2F(\theta^*) + 2(\rho - \gamma) F'(\theta^*) \frac{d\theta^*}{d\rho} = 2F(\theta^*) + 2(\rho - \gamma) F'(\theta^*) \frac{1}{2F'(\theta^*)\gamma - 1}.$$

We claim this is positive. This is equivalent to showing

$$F(\theta^*) > \frac{(\rho - \gamma) F'(\theta^*)}{1 - 2F'(\theta^*)\gamma}$$

which is the same as

$$F(\theta^*) > (\rho - \gamma + 2\gamma F(\theta^*))\, F'(\theta^*) = (2\rho + \theta^*)\, F'(\theta^*) \tag{34}$$

where the equality uses (15). Because $2F'(\theta^*)\gamma < 1$,

$$(2\rho + \theta^*)\, F'(\theta^*) < (2\rho + \theta^*)\, /2\gamma.$$

Therefore, to show (34) holds we need to show that $2F(\theta^*)\gamma > 2\rho + \theta^*$. This is true because of (15) and $\rho < \gamma$. ■