Implicit Incentives and Delegation in Teams: Online Appendix

Harvey Upton*

A: Expansion of the Contracting Space

In this appendix, we consider expansions of the contracting space that allow wage payments to be conditioned on additional elements beyond project success. This can be as part of a formal contract (i.e. in environments where the state and adopted production method are verifiable to third parties) or as part of a relational agreement. We start by considering an environment of **perfect commitment**, where the principal can credibly commit to any contract with payments that are conditional on these variables. This corresponds to the case of a formally contractable state and production method, but also to the case of a relational contract in which the principal has no difficulty in committing to the informal agreement, for instance if her discount factor is sufficiently high. We subsequently use these results to study an environment of **limited commitment**, where the variables cannot form part of a formal contract and the principal has difficulty in credibly committing to certain wage schemes; in this case, the wage schedule must be adjusted in order to remain self-enforcing under an informal agreement.

Perfect Commitment

Suppose the principal can perfectly commit to wages that condition on the state and selected production method. As in the main text, for any given punishment equilibrium and under either organisational structure, the principal cannot benefit from positive wage payments following an unsuccessful project, since this would increase wage expenditure while undermining effort incentives. Accordingly, she designs a wage schedule $w_i(\gamma_j, \omega_k) : \{\gamma_1, \gamma_2\} \times \{\omega_1, \omega_2\} \rightarrow \mathbb{R}_+$ for $i \in \{A, B\}$, which specifies payments following a successful project, conditional on the state and production method. Her goal is to minimise total expected wage costs:

$$p_2 \left[r \left(w_A(\gamma_1, \omega_1) + w_B(\gamma_1, \omega_1) \right) + (1 - r) \left(w_A(\gamma_2, \omega_2) + w_B(\gamma_2, \omega_2) \right) \right] \tag{1}$$

while implementing her desired outcome as a subgame-perfect equilibrium; that is, high effort from both agents and the correct production method in each period.

Centralisation

In each period under centralisation, the principal observes the state and implements the correct production method. Accordingly, the wages $w_i(\gamma_1, \omega_2)$ and $w_i(\gamma_2, \omega_1)$ can never be paid and, as such, their value is irrelevant to the design problem. Moreover, the agents play a symmetric role in the production process. For Agent i, the principal chooses the wage payments $w_i(\gamma_1, \omega_1)$ and $w_i(\gamma_2, \omega_2)$ which minimise costs while inducing high effort. As discussed in the main text, the principal cannot benefit from designing the wage

*Humboldt-Universität zu Berlin. Contact: harvey.upton@hu-berlin.de

scheme such that either agent is induced to work during the punishment phase of the game. Accordingly, on the punishment path following a deviation by either agent, both agents shirk repeatedly.

The expected utility in each period for Agent i when the team adheres to the principal's desired outcome (high effort from both) is:

$$\widehat{U}_i = r\left[p_2 w_i(\gamma_1, \omega_1) - c_i^1\right] + (1-r)\left[p_2 w_i(\gamma_2, \omega_2) - c_i^2\right] - e \tag{2}$$

while expected utility on the punishment path when both agents repeatedly shirk is:

$$\widehat{P}_i^{cent} = r\left[p_0 w_i\left(\gamma_1, \omega_1\right) - c_i^1\right] + (1-r)\left[p_0 w_i\left(\gamma_2, \omega_2\right) - c_i^2\right] \tag{3}$$

It follows that the incentive compatibility constraints for Agent i, for states ω_1 and ω_2 respectively, can be expressed as:

$$(1-\delta)(p_2 w_i\left(\gamma_1, \omega_1\right) - e - c_i^1) + \delta\widehat{U}_i \geq (1-\delta)(p_1 w_i\left(\gamma_1, \omega_1\right) - c_i^1) + \delta\widehat{P}_i^{cent} \tag{4}$$

and:

$$(1-\delta)(p_2 w_i\left(\gamma_2, \omega_2\right) - e - c_i^2) + \delta\widehat{U}_i \geq (1-\delta)(p_1 w_i\left(\gamma_2, \omega_2\right) - c_i^2) + \delta\widehat{P}_i^{cent} \tag{5}$$

The principal's problem is then to minimise (1), subject to these constraints and the restriction that all wage payments must be non-negative.

Proposition A1. *Under centralisation, the principal minimises the costs of implementing her desired outcome as a subgame-perfect equilibrium by setting:*

$$w_i\left(\gamma_1, \omega_1\right) = w_i\left(\gamma_2, \omega_2\right) = w^{cent} = \frac{e}{p_2 - (1-\delta)p_1 - \delta p_0} \tag{6}$$

with $w_i\left(\gamma_1, \omega_2\right)$ and $w_i\left(\gamma_2, \omega_1\right)$ set arbitrarily, for $i \in \{A, B\}$.

Intuitively, under centralisation the principal selects the correct production method in each period, so that wages need only to induce high effort from both agents, given the threat of repeated shirking on the punishment path. Since the agents cannot influence the choice of production method through their behaviour, wages cannot be reduced by conditioning on either the implemented method or the state. It then follows that the optimal wage structure is identical to that of the environment considered in the main text, whereby agents are rewarded with the bonus payment w^{cent} if and only if the project is successful. As such, the principal cannot benefit from the increased information available with an expanded contracting space.

Delegation

Under delegation, the production method in each period is selected by Agent A, so that the wage scheme must provide incentives for high effort and for the correct production method to be chosen. Accordingly, the payments $w_i(\gamma_1, \omega_2)$ and $w_i(\gamma_2, \omega_1)$ may now play a role in determining the principal's expected wage costs. As in the main text, the key benefit of delegation is that the agents' increased discretion over the production process allows for alternative punishment play following a deviation, compared to centralisation. As such, we restrict attention to the case where (γ_1 *and shirk*; *shirk*) is played in both states on the punishment path;

this allows for a direct comparison with the wage scheme derived in Proposition 3, in order to study the implications of the expanded contracting space.

To begin, note that the principal can never be adversely affected by an expansion of the contracting space. In particular, she always has the option of ignoring the additional information now available to her and instead conditioning wages only on the project outcome. This would induce the same stage games as in the main text, Γ_1^{del} and Γ_2^{del}, so that all subsequent results apply. In other words, the principal can always replicate the wage scheme and hence expected wage costs from the environment studied in the main text. Combined with Proposition A1, this immediately implies that if delegation is optimal when the principal cannot condition wages on the state and production method, it will also be optimal when she can. In the remainder of this section, we show that in addition, there exist cases where the expanded contracting space allows her to strictly reduce wage costs. Accordingly, there are situations where the ability to contract on the state and production method cause delegation to become optimal as an organisational design.

As in the case of centralisation, Agent i's expected utility in each period when the principal's desired outcome is played is given by (2). During the punishment phase of the game, when (γ_1 *and shirk*; *shirk*) is played regardless of the state, the expected per-period utility of Agent i is:

$$\widehat{P}_i^{del} = rp_0 w_i(\gamma_1, \omega_1) + (1-r)\, q_0 w_i(\gamma_1, \omega_2) - c_i^1 \tag{7}$$

Since the agents play disparate roles in the production process under delegation, we consider the incentive compatibility constraints for each in turn. For Agent A, in state ω_1 we require that:

$$(1-\delta)\left(p_2 w_A(\gamma_1, \omega_1) - e - c_A^1\right) + \delta \widehat{U}_A \geq (1-\delta)\left(p_1 w_A(\gamma_1, \omega_1) - c_A^1\right) + \delta \widehat{P}_A^{del} \tag{8}$$

$$(1-\delta)\left(p_2 w_A(\gamma_1, \omega_1) - e - c_A^1\right) + \delta \widehat{U}_A \geq (1-\delta)\left(q_2 w_A(\gamma_2, \omega_1) - e - c_A^2\right) + \delta \widehat{P}_A^{del} \tag{9}$$

$$(1-\delta)\left(p_2 w_A(\gamma_1, \omega_1) - e - c_A^1\right) + \delta \widehat{U}_A \geq (1-\delta)\left(q_1 w_A(\gamma_2, \omega_1) - c_A^2\right) + \delta \widehat{P}_A^{del} \tag{10}$$

while in state ω_2 we require:

$$(1-\delta)\left(p_2 w_A(\gamma_2, \omega_2) - e - c_A^2\right) + \delta \widehat{U}_A \geq (1-\delta)\left(p_1 w_A(\gamma_2, \omega_2) - c_A^2\right) + \delta \widehat{P}_A^{del} \tag{11}$$

$$(1-\delta)\left(p_2 w_A(\gamma_2, \omega_2) - e - c_A^2\right) + \delta \widehat{U}_A \geq (1-\delta)\left(q_2 w_A(\gamma_1, \omega_2) - e - c_A^1\right) + \delta \widehat{P}_A^{del} \tag{12}$$

$$(1-\delta)\left(p_2 w_A(\gamma_2, \omega_2) - e - c_A^2\right) + \delta \widehat{U}_A \geq (1-\delta)\left(q_1 w_A(\gamma_1, \omega_2) - c_A^1\right) + \delta \widehat{P}_A^{del} \tag{13}$$

Similarly, for Agent B, the respective incentive compatibility constraints associated with states ω_1 and ω_2 are:

$$(1-\delta)\left(p_2 w_B(\gamma_1, \omega_1) - e - c_B^1\right) + \delta \widehat{U}_B \geq (1-\delta)\left(p_1 w_B(\gamma_1, \omega_1) - c_B^1\right) + \delta \widehat{P}_B^{del} \tag{14}$$

$$(1-\delta)\left(p_2 w_B(\gamma_2, \omega_2) - e - c_B^2\right) + \delta \widehat{U}_B \geq (1-\delta)\left(p_1 w_B(\gamma_2, \omega_2) - c_B^2\right) + \delta \widehat{P}_B^{del} \tag{15}$$

As in the case of centralisation, the principal's expected wage costs for a particular wage schedule are given by (1). Her goal is to minimise these costs, while designing a wage scheme which implements her desired outcome as subgame-perfect equilibrium.

Proposition A2. *Suppose we have:*

$$\Delta_A \in \left[\frac{p_0 e}{p_2-(1-\delta)p_1-\delta p_0}, \frac{p_1 e}{p_2-(1-\delta r)\,p_1-\delta r p_0}\right] \tag{16}$$

and

$$\Delta_B \leq \frac{e\,[p_2-p_1-(1-\delta)p_1+p_0\,(1-\delta r)]}{(p_1-p_0)\,\delta\,(1-r)} \tag{17}$$

Then a wage scheme that sets:

$$w_i(\gamma_1,\omega_1) = w_i(\gamma_2,\omega_2) = \frac{e+\delta\,(1-r)\,\Delta_i}{p_2-(1-\delta)p_1-\delta r p_0} \tag{18}$$

for $i \in \{A,B\}$ *and all other wages equal to zero implements the principal's desired outcome as a subgame-perfect equilibrium. The principal's wage payments to Agents A and B are strictly lower than* w_A^{del} *and* w_B^{del}, *respectively, and, for any set of remaining model parameters, the interval* (16) *is non-empty.*

Proposition A2 shows that there exist cases where the additional information that can be incorporated into the wage scheme under an expanded contracting space strictly reduces the principal's expected costs. Intuitively, setting all wages equal to zero following adoption of an incorrect production method has two distinct effects on the agents' incentive compatibility constraints. First, Agent A is less tempted to deviate from the principal's desired outcome by selecting his favoured production method, since this implies a zero wage with certainty. Second, following a deviation, the expected per-period utility on the punishment path, (7), is reduced for both agents, since they receive a lower expected wage. Both effects slacken incentive compatibility constraints, allowing for a strict reduction in total expected wage payments to both agents relative to the case studied in the main text.[1]

Limited Commitment

In this section, we use the foregoing results to study the case where the state and production method can only be incorporated into the wage scheme as part of a relational contract. We denote the principal's discount factor by δ_P. As discussed in the foregoing, when δ_P is sufficiently high, the principal can perfectly commit to wage schemes that condition on this additional information, so that the results in the previous section apply. However, when δ_P is low, the wage scheme must be adjusted in order to remain self-enforcing.

Specifically, for a given combination of production method and state, the principal outlines a wage to be paid following a successful project. However, since the contract is informal, she also has the option of reneging on this promise by instead paying the lowest possible bonus outlined by the formal contract; by doing so, she becomes untrustworthy to the agents and thus sacrifices the ability to make use of informal incentives in future periods.[2] From a technical perspective, the case of relational contracting with limited

[1] Note that, as with Proposition 3, Proposition A2 does not state that this wage scheme is optimal for the principal. In particular, for a given set of model parameters, she may be able to reduce her costs by designing a wage scheme which implements an alternative punishment equilibrium. For instance, by setting $w_i(\gamma_1,\omega_1) = w_i(\gamma_2,\omega_2) = w^{cent}$ and $w_i(\gamma_1,\omega_2) = w_i(\gamma_2,\omega_1) = 0$ for both agents, the principal can achieve identical wage costs to the case of centralisation so long as Δ_A is not too high, analogous to Proposition 2. Moreover, the expanded contracting space may allow for equilibria which are not feasible when the principal cannot condition wages on this additional information.

[2] Note that since project success is verifiable, the principal cannot deviate to the (zero) wage paid following a failed project. Moreover, since wages following project failure are zero, there is no incentive to deviate after an unsuccessful project.

commitment places additional constraints on the principal's design of the wage structure, relative to the environment considered in the previous section.

To investigate the case of limited commitment, we make use of the foregoing results. Under centralisation, Proposition A1 establishes that the principal cannot benefit from incorporating additional information into the wage scheme, even when she can perfectly commit to doing so. Accordingly, it is optimal to set all wages following a successful project equal to w^{cent}, so that informal payments play no role. Under delegation, Proposition A2 shows that there exist cases where the principal can strictly benefit from conditioning wages on the production method and state, in particular by setting wages following adoption of the incorrect production method equal to zero. This is potentially problematic when contracts are relational, since the principal has a large short-term benefit from a deviation.

As in the previous section, in order to be incentive compatible given repeated play of (γ_1 *and shirk*; *shirk*) in both states on the punishment path, the wage schedule under delegation must satisfy the constraints (9)-(15). However, the requirement that the relational contract is self enforcing also imposes additional constraints on the wage scheme. Let us denote the principal's expected per-period payoff when the agents are playing her desired outcome as:

$$U_P := p_2 \left[R - r\left(w_A(\gamma_1,\omega_1) + w_B(\gamma_1,\omega_1)\right) - (1-r)\left(w_A(\gamma_2,\omega_2) + w_B(\gamma_2,\omega_2)\right)\right] \tag{19}$$

and let $\widehat{P}_P$ denote her payoff after she has deviated from a promise to pay a prescribed wage and thus becomes untrustworthy to the agents.[3] Moreover, for a particular wage scheme we define:

$$\boldsymbol{w}^{min} = \min_{j\in\{1,2\},k\in\{1,2\}} w_A(\gamma_j,\omega_k) + \min_{j\in\{1,2\},k\in\{1,2\}} w_B(\gamma_j,\omega_k) \tag{20}$$

as the smallest possible total wage expenditure that the principal must bear following a successful project, according to the formal contract. The principal's feasibility constraints in states ω_1 and ω_2 can then be written as:

$$(1-\delta_P)\left(R - w_A(\gamma_1,\omega_1) - w_B(\gamma_1,\omega_1)\right) + \delta_P U_P \geq (1-\delta_P)\left(R - \boldsymbol{w}^{min}\right) + \delta_P \widehat{P}_P \tag{21}$$

$$(1-\delta_P)\left(R - w_A(\gamma_2,\omega_2) - w_B(\gamma_2,\omega_2)\right) + \delta_P U_P \geq (1-\delta_P)\left(R - \boldsymbol{w}^{min}\right) + \delta_P \widehat{P}_P \tag{22}$$

The two constraints can be explained similarly. After a successful project where the correct production method has been used, in both states the principal faces a choice between paying the prescribed wages to the agents, or reneging on the relational contract and paying them both the smallest viable wage.[4] In the former case, the relational contract remains intact and the principal receives the expected utility U_P in each subsequent period; in the latter case, relational contracting breaks down and she receives $\widehat{P}_P$ thereafter. Defining:

$$\boldsymbol{w}^{max} := \max_{j\in\{1,2\}} w_A(\gamma_j,\omega_j) + w_B(\gamma_j,\omega_j) \tag{23}$$

[3] The principal's fallback utility $\widehat{P}_P$ is determined by the parties' behaviour after a deviation from the relational contract. There are several possibilities here. Some papers assume that the parties' interaction continues under spot employment (Baker et al., 2002, Kvaløy and Olsen, 2006) or under the optimal formal contract in the absence of relational incentives (Baker et al., 1994), while others assume that the relationship breaks down, with all parties subsequently receiving their outside options (Levin, 2003). We do not examine this issue in detail.

[4] Note that since the principal reveals herself to be untrustworthy to both agents following a deviation, it is optimal for her to pay them both their respective minimum payments.

allows us to combine (21) and (22); rearranging then yields:

$$\frac{\delta_P}{1-\delta_P}\left(U_P - \widehat{P}_P\right) \geq \boldsymbol{w}^{max} - \boldsymbol{w}^{min} \tag{24}$$

This condition is familiar from the relational contracting literature and states that the difference between the principal's maximum and minimum possible wage payments cannot exceed the discounted future benefit that she accrues from her ability to make use of informal agreements.

As discussed, for sufficiently high values of δ_P, (24) will not bind. This is also true when there are substantial costs of reneging on the relational agreement; for instance, if R is very large and the principal suffers a loss in expected output following a deviation (e.g. due to shirking by agents or a breakdown of the relationship). In these cases, the principal can continue to implement the wage scheme derived in Proposition A2, even when contracts must be relational.

However, for smaller values of δ_P this payment schedule is not feasible, since the principal is tempted to deviate by paying the zero wages associated with adoption of the incorrect production method. Accordingly, once (24) binds, incentives must be adjusted so that the difference in possible wages following project success is sufficiently small; further reductions in δ_P therefore require increases in (at least some of) the payments $w_i(\gamma_1, \omega_2)$ and $w_i(\gamma_2, \omega_1)$ for $i \in \{A, B\}$, so that the principal's profit is strictly reduced. In the limit, as $\delta_P \to 0$, we require that

$$\boldsymbol{w}^{max} - \boldsymbol{w}^{min} = 0 \tag{25}$$

so that the optimal contract pays the same bonus following project success regardless of production method and state, similar to the case considered in the main text.[5] In this respect, one can think of relational contracting with limited commitment as representing an intermediate case between perfect commitment as presented in the previous section of this appendix and the environment of the main text; in particular, the principal's wage costs under relational contracting with limited commitment are bounded by these two cases.

Discussion

This appendix considers expansions of the contracting space, relative to the environment of the main text. We show that the ability to condition wages on the production method and state does not affect the principal's wage costs under centralisation, but weakly reduces them under delegation. These results hold whether she can perfectly commit to a wage scheme that incorporates this additional information or whether this must be done using an informal agreement. It follows that the expanded contracting space can only favour the use of delegation.

We also show that the key results derived in the main text continue to hold in this environment. First, delegation expands the potential scope of peer sanction, allowing for alternative punishment play, which can generate increased implicit incentives relative to the case of centralisation. Second, from Proposition A2, for a given pair (Δ_A, Δ_B) that satisfy the conditions (16) and (17), a reduction in Δ_B strictly reduces the principal's wage costs while continuing to allow the principal's desired outcome to be achieved as a subgame-perfect equilibrium. In particular, this implies that delegation is especially attractive for the principal with a diverse team composition ($\Delta_B < 0$), similar to the main text.[6]

[5] Note that since the wages $w_i(\gamma_1, \omega_2)$ and $w_i(\gamma_2, \omega_1)$ for $i \in \{A, B\}$ are never paid on the equilibrium path, they are not subject to this restriction. However, setting these payments to be strictly greater than $w_i(\gamma_1, \omega_1)$ and $w_i(\gamma_2, \omega_2)$ can never be optimal, since this serves to tighten the incentive compatibility constraints (9)-(15).

[6] We do not provide here an exhaustive analysis of all the paper's results in this setting; nonetheless, many of them can be shown to be qualitatively similar to the case studied in the main text. A small number can be shown to no longer apply in

Proof of Proposition A1. We can rewrite (4) and (5) as:

$$(1-\delta)p_2 w_i(\gamma_1,\omega_1)+\delta\left[rp_2 w_i(\gamma_1,\omega_1)+(1-r)\,p_2 w_i(\gamma_2,\omega_2)\right]-e \geq (1-\delta)p_1 w_i(\gamma_1,\omega_1)+\delta\left[rp_0 w_i(\gamma_1,\omega_1)+(1-r)\,p_0 w_i(\gamma_2,\omega_2)\right] \quad (26)$$

and:

$$(1-\delta)p_2 w_i(\gamma_2,\omega_2)+\delta\left[rp_2 w_i(\gamma_1,\omega_1)+(1-r)\,p_2 w_i(\gamma_2,\omega_2)\right]-e \geq (1-\delta)(p_1 w_i(\gamma_2,\omega_2))+\delta\left[rp_0 w_i(\gamma_1,\omega_1)+(1-r)\,p_0 w_i(\gamma_2,\omega_2)\right] \quad (27)$$

At least one of these constraints must bind, otherwise the principal would pay a constant wage of zero, which cannot create effort incentives. Suppose that (26) binds; rearranging then yields:

$$w_i(\gamma_1,\omega_1)=\frac{e-\delta(1-r)\left[p_2-p_0\right]w_i(\gamma_2,\omega_2)}{\left[(1-\delta)(p_2-p_1)+\delta r(p_2-p_0)\right]} \quad (28)$$

so that the principal's objective function can be rewritten as:

$$\min_{w_i(\gamma_2,\omega_2)} \frac{rp_2 e+(1-r)\,p_2 w_i(\gamma_2,\omega_2)(1-\delta)(p_2-p_1)}{\left[(1-\delta)(p_2-p_1)+\delta r(p_2-p_0)\right]} \quad (29)$$

Clearly, this is increasing in $w_i(\gamma_2,\omega_2)$. Accordingly, if (27) does not bind, the principal would set $w_i(\gamma_2,\omega_2)=0$. However, inserting this into (27) along with (28), rearranging and simplifying yields:

$$-\frac{(1-\delta)e(p_2-p_1)}{(1-\delta)(p_2-p_1)+\delta r(p_2-p_0)}\geq 0 \quad (30)$$

which is clearly not satisfied; accordingly, it cannot be the case that (26) binds while (27) does not. A similar argument implies that we cannot have (27) binding with (26) slack, so that both (26) and (27) must bind at the optimal solution. Solving these equations simultaneously then yields (6).

Since this wage scheme is identical to the one derived in the case of centralisation in the main text, it follows that the stage games are also identical to that case. Accordingly, both agents shirking in each period is a Nash Equilibrium of the stage game in each state, so that this punishment path is feasible and the principal's desired outcome is supported as a subgame-perfect equilibrium. □

Proof of Proposition A2. We start with Agent A. First, for the wage scheme outlined by the proposition, it is straightforward to verify that (10) and (13) imply (9) and (12), respectively. Moreover, since $w_A(\gamma_1,\omega_1)=w_A(\gamma_2,\omega_2)$, (8) and (11) are equivalent, while (13) implies (10). Altogether, in order to guarantee that (8)-(13) are satisfied, we require that:

$$(1-\delta)\left((p_2-p_1)\frac{e+\delta(1-r)\Delta_A}{p_2-(1-\delta)p_1-\delta rp_0}-e\right)+\delta\left(\widehat{U}_A-\widehat{P}_A^{del}\right)\geq 0 \quad (31)$$

$$(1-\delta)\left(p_2\frac{e+\delta(1-r)\Delta_A}{p_2-(1-\delta)p_1-\delta rp_0}-e-\Delta_A\right)+\delta\left(\widehat{U}_A-\widehat{P}_A^{del}\right)\geq 0 \quad (32)$$

general; for instance, the wage payments (18) are no longer affected by changes in κ, in contrast to Proposition 4.

Using (2) and (7), (31) can be shown to hold with equality, while (32) is satisfied *iff.*:

$$\frac{ep_1}{p_2 - (1-\delta r)\,p_1 - \delta r p_0} \geq \Delta_A \tag{33}$$

which is guaranteed by (16). For Agent B, for the wage scheme outlined by the proposition, (14) and (15) become equivalent and are satisfied with equality. Accordingly, the wage scheme satisfies all incentive compatibility constraints.

Next, in order for (γ_1 *and shirk*; *shirk*) to be Nash Equilibria of the stage games in both states, the following conditions must hold for Agent A:

$$p_0 w_A(\gamma_1,\omega_1) - c_A^1 \geq p_1 w_A(\gamma_1,\omega_1) - e - c_A^1 \tag{34}$$

$$p_0 w_A(\gamma_1,\omega_1) - c_A^1 \geq q_1 w_A(\gamma_2,\omega_1) - e - c_A^2 \tag{35}$$

$$p_0 w_A(\gamma_1,\omega_1) - c_A^1 \geq q_0 w_A(\gamma_2,\omega_1) - c_A^2 \tag{36}$$

$$q_0 w_A(\gamma_1,\omega_2) - c_A^1 \geq q_1 w_A(\gamma_1,\omega_2) - e - c_A^1 \tag{37}$$

$$q_0 w_A(\gamma_1,\omega_2) - c_A^1 \geq p_1 w_A(\gamma_2,\omega_2) - e - c_A^2 \tag{38}$$

$$q_0 w_A(\gamma_1,\omega_2) - c_A^1 \geq p_0 w_A(\gamma_2,\omega_2) - c_A^2 \tag{39}$$

(35) and (36) are trivially satisfied for any $w_A(\gamma_1,\omega_1) \geq 0$, while (37) is satisfied for $w_A(\gamma_1,\omega_2) = 0$. Using the wage schedule outlined by the proposition, (34), (38) and (39) can be respectively expressed as:

$$\frac{e\left[p_2 - p_1 - (1-\delta)p_1 + p_0\,(1-\delta r)\right]}{(p_1 - p_0)\,\delta(1-r)} \geq \Delta_A \tag{40}$$

$$\Delta_A \geq \frac{e\left[-p_2 + (2-\delta)p_1 + \delta r p_0\right]}{p_2 - [1-\delta r]\,p_1 - \delta r p_0} \tag{41}$$

$$\Delta_A \geq \frac{p_0 e}{p_2 - (1-\delta)p_1 - \delta p_0} \tag{42}$$

Note that:

$$\frac{e\left[p_2 - p_1 - (1-\delta)p_1 + p_0\,(1-\delta r)\right]}{(p_1 - p_0)\,\delta(1-r)} \geq \frac{ep_1}{p_2 - [1-\delta r]\,p_1 - \delta r p_0} \tag{43}$$

$$\iff \left[p_2 - (1-\delta)\,p_1 - \delta r p_0\right]\left[p_2 - p_1 - (1-\delta r)\,(p_1 - p_0)\right] \geq 0 \tag{44}$$

which is satisfied since both of the square brackets in (44) are positive. Similarly,

$$\frac{p_0 e}{p_2 - (1-\delta)p_1 - \delta p_0} \geq \frac{e\left[-p_2 + (2-\delta)p_1 + \delta r p_0\right]}{p_2 - [1-\delta r]\,p_1 - \delta r p_0} \tag{45}$$

$$\iff \left[p_2 - (1-\delta)p_1 - r\delta p_0\right]\left[p_2 - p_1 - (1-\delta)\,(p_1 - p_0)\right] \geq 0 \tag{46}$$

which is also satisfied. It follows that (40)-(42) are satisfied by any Δ_A in the interval (16).

For Agent B we require that:

$$p_0 w_B(\gamma_1, \omega_1) - c_B^1 \geq p_1 w_B(\gamma_1, \omega_1) - e - c_B^1 \quad (47)$$

$$q_0 w_B(\gamma_1, \omega_2) - c_B^1 \geq q_1 w_B(\gamma_1, \omega_2) - e - c_B^1 \quad (48)$$

For the wage schedule outlined by the proposition, (48) holds trivially, while (47) becomes equivalent to (17). Moreover, it is straightforward to verify that the LHS of (17) is strictly positive. Altogether, all constraints are satisfied, so that (γ_1 *and shirk*; *shirk*) is a Nash Equilibrium of the stage game in both states of the world. Along with the incentive compatibility constraints, this implies that the wage scheme outlined by the proposition implements the principal's desired outcome as a subgame-perfect equilibrium. It is straightforward to verify that the wages to each agent are reduced relative to the case studied in the main text, while

$$\frac{ep_1}{p_2 - [1 - \delta r]\, p_1 - \delta r p_0} > \frac{p_0 e}{p_2 - (1 - \delta) p_1 - \delta p_0} \quad (49)$$

$$\iff p_2 - (1 - \delta) p_1 - \delta r p_0 > 0 \quad (50)$$

so that the interval (16) always exists and is non-empty. □

B: Collusion by the Agents

In the main text, our analysis focused on whether a particular combination of organisational structure and wage scheme could implement the principal's desired outcome as a subgame-perfect equilibrium. However, we did not consider whether such equilibria are optimal from the agents' point of view. It is possible that for a particular organisational design and wage schedule, there exist alternative equilibria that allow for a Pareto improvement of the agents' expected payoffs. In this appendix, we study whether such collusion is possible for the agents, considering each organisational structure in turn.[7]

Under centralisation, the agents cannot control the choice of production method and thus can only collude with respect to effort provision. As shown by Che and Yoo (2001), when this is the case, a wage structure of $\{w^{cent}, w^{cent}\}$ implies that the agents' surplus is maximised in each period when they both undertake high effort.[8] It follows that no other set of strategies, regardless of whether they form an equilibrium, can improve on repeated play of $(work, work)$. Intuitively, implicit incentives are created in the repeated game by the threat of shirking in future periods. For this threat to be effective, the agents must be better off when they both work compared to the case where they both shirk, so that the surplus is higher. Moreover, supermodularity of the production function implies that the surplus cannot be highest when only one agent works.

Next, we study collusion when the production method choice is delegated to Agent A. Since for $\Delta_A < \widehat{\Delta}_A$ delegation has identical wage costs to centralisation, we focus on the case of $\Delta_A \in \left(\widehat{\Delta}_A, \Delta_A^{max}\right)$ with $\Delta_B < \Delta_B^{max}$ and in particular consider the following question: can the principal continue to achieve strictly lower wage costs through delegation when the equilibrium is additionally required to be collusion-proof?

Studying collusive outcomes under delegation is significantly more complicated than under centralisation for a number of reasons. First, Agent A's control over the production method implies a wider range of potential outcomes that could be implemented as part of a collusive agreement. For instance, the two agents may agree to always adopt a particular production method, regardless of the state, if they have particularly strong preferences for such an outcome, as well as coordinating on whether to provide high or low effort. Second, control over the production method also implies that agents' collusive strategies will likely condition on the state. Third, from a technical perspective, the asymmetric wage scheme associated with delegation further complicates analysis, since the agents may have differing preferences over the relative attractiveness of working and shirking given a particular production method and state. Nonetheless, we can show the following result.

Proposition B1. *There always exists a non-empty set of* (Δ_A, Δ_B) *combinations, where* $\Delta_A \in \left(\widehat{\Delta}_A, \Delta_A^{max}\right)$ *and* $\Delta_B < \Delta_B^{max}$*, such that (i) under delegation, the wage scheme* $\left\{w_A^{del}, w_B^{del}\right\}$ *implements the principal's desired outcome at strictly lower costs than centralisation and (ii) the resulting equilibrium is collusion-proof.*

The proof of Proposition B1 shows that there always exists a set of (Δ_A, Δ_B) combinations such that, under delegation, the principal's desired outcome also maximises the agents' surplus, similar to the case of centralisation. Specifically, we consider the agents' surplus as a function of $\Delta_{sum} = \Delta_A + \Delta_B$, which

[7] To be clear, our notion of collusion-proofness requires that there exists no other subgame-perfect equilibrium which is Pareto dominant for the agents, in the sense that it offers a weakly better expected payoff to both and a strictly better payoff to at least one. Other notions of collusion-proofness have been considered in the literature; for instance, Che and Yoo's (2001) concept of 'team equilibrium' requires that an outcome is the subgame-perfect equilibrium which yields the highest total surplus to the agents, while Glover and Kim (2023) primarily study 'regret-free collusion'. As discussed in the model setup, we rule out the possibility of explicit side-contracting between agents à la Itoh (1992, 1993), which can be thought of as an alternative form of collusion.

[8] We omit a formal proof, instead referring the reader to Proposition 4 and the related discussion in their paper.

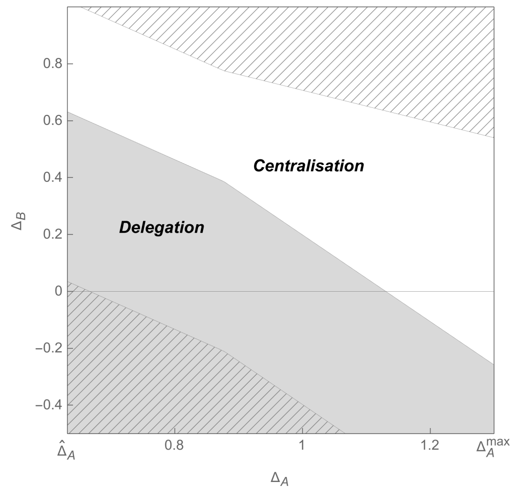

Figure 1: The possibility for collusion.

captures their aggregate preferences over production methods. We show that there always exists a non-empty interval such that, for all values of Δ_{sum} within this interval, any deviation away from the principal's desired outcome leads to a reduction in the agents' joint expected payoff, so that they have no incentives to deviate from the prescribed equilibrium under delegation. However, we also show that this equilibrium may fail to be collusion-proof when Δ_{sum} is either too large or too small.

Figure 1 reproduces the example studied in Figure 4 in the main text, but also displays the implications of changes in Δ_A and Δ_B for the agents' surplus and the possibility of collusion.[9] As before, the white and grey areas respectively show the range of values over which centralisation and delegation are optimal for the principal. The two new shaded areas show the regions where, under delegation, the principal's desired outcome does not maximise the agents' surplus.

First, when Δ_{sum} is very large, the agents have strong aggregate preferences for adopting production method γ_1. In this case, agents may be able to increase their joint surplus by agreeing to implement production method γ_1 in both states of the world; in Figure 1, this possibility is illustrated by the shaded region in the top-right of the graphic. However, from Proposition 4, the wage w_i^{del} is increasing in Δ_i for $i \in \{A, B\}$; accordingly, loosely speaking, we can say that the total wage payment under delegation is increasing in Δ_{sum}. Moreover, the proof of Proposition B1 shows that in cases where Δ_{sum} is sufficiently large such that the aforementioned collusion is a concern, the principal always faces lower costs under centralisation than delegation. Put differently, as long as delegation is the optimal organisational design for the principal, Δ_{sum} can never become sufficiently high such that collusion is a concern. Corresponding to this result, in Figure 1 the shaded region in the top-right does not intersect with the grey area in which delegation is optimal.

Second, there is a shaded area in the bottom-left of the graphic where Δ_{sum} becomes very low. This is problematic for the principal since, for the aforementioned reason, total wages under delegation will typically be lower than those under centralisation and thus the principal would prefer to adopt delegation. Collusion in this region may occur for two distinct reasons. First, as discussed in the foregoing, total wages are decreasing as Δ_{sum} becomes lower. Accordingly, for sufficiently small Δ_{sum}, the agents' wages can become so low that working is no longer worthwhile, in which case their joint surplus is maximised when both agents shirk. Second, if Δ_A and Δ_B are so low that Δ_{sum} is sufficiently negative, the agents have strong aggregate

[9]For the parameters associated with this specific example, see the discussion in the main text.

preferences for adopting γ_2 in both states, which may also allow for an increased surplus by deviating from the principal's desired outcome.

Proposition B1 therefore establishes that there exists the *potential* for collusion; i.e. that there are alternative outcomes which *could* increase the agents' joint surplus. Collusion only becomes feasible if, in addition, these outcomes can be sustained as a subgame-perfect equilibrium, given that the principal selects delegation and offers the wage scheme $\{w_A^{del}, w_B^{del}\}$. However, the large literature which analyses attainable outcomes in repeated games (in particular by providing Folk Theorems; see for instance Fudenberg and Maskin, 1986) emphasises that the set of equilibria will typically be large, so long as the players are sufficiently patient.[10] Accordingly, we can expect there to be many cases in which there exist alternative equilibria that yield a higher joint surplus to the agents than the principal's desired outcome.

For the reasons discussed earlier in this appendix, characterising the agents' optimal collusive equilibrium for each set of parameters is complex, and is beyond the scope of our analysis. Instead, we present an example of a collusive deviation from the principal's desired outcome under delegation; that is, a situation where there exists an alternative outcome which (i) is weakly better for both players and strictly better for one and (ii) can be sustained as an equilibrium.

Consider the following set of parameters: $e = 1$, $r = \frac{1}{2}$, $\delta = \frac{1}{2}$; we let $q_l = \kappa p_l$ for $l \in \{0, 1, 2\}$, with $p_2 = \frac{1}{2}$, $p_1 = \frac{502}{1024}$, $p_0 = \frac{497}{1024}$ and $\kappa = \frac{1}{2}$. We set $c_A^1 = 0$ and $c_A^2 = \frac{2000}{95}$ so that $\Delta_A = \frac{2000}{95} \approx 21.05$. We also set $c_B^1 = c_B^2 = 1$, so that $\Delta_B = 0$. Note that with these parameters, we have $\widehat{\Delta}_A = 19.88$, $\Delta_A^{max} = 25.35$ and $\Delta_B^{max} = 25.85$; accordingly, we have $\Delta_A \in \left(\widehat{\Delta}_A, \Delta_A^{max}\right)$ and $\Delta_B < \Delta_B^{max}$ as required by Proposition 3. In addition, $\Delta_A^{crit} = 21.14$ so that we also have $\Delta_A < \Delta_A^{crit}$.

The total wage payment under centralisation is given by $2w^{cent} = 163.84$. Moreover, the expected utility of each agent is as follows:

$$EU_A^{cent} = p_2 w^{cent} - e - r c_A^1 - (1-r) c_A^2 = 29.43 \tag{51}$$

$$EU_B^{cent} = p_2 w^{cent} - e - r c_B^1 - (1-r) c_B^2 = 38.96 \tag{52}$$

so that both agents prefer to participate.

Keeping in mind that $\Delta_A < \Delta_A^{crit}$, the sum of the wages under delegation is $w_A^{del} + w_B^{del} = 99.66$. Their respective expected utilities are given by:

$$EU_A^{del} = p_2 w_A^{del} - e - r c_A^1 - (1-r) c_A^2 = 31.45 \tag{53}$$

$$EU_B^{del} = p_2 w_B^{del} - e - r c_B^1 - (1-r) c_B^2 = 4.86 \tag{54}$$

so that again, both agents prefer to participate. It follows that, in the absence of collusion, the principal strictly prefers to implement delegation.

We now show that if the principal chooses to implement delegation, there exists a collusive strategy which yields strictly higher utility to both players. Suppose that in state ω_1, Agent A always plays γ_1 *and shirk* and in state ω_2, always plays γ_2 *and shirk*. That is, Agent A always implements the correct method of production

[10] There is however a small, but important, difference to these setups, since in the current framework the agent's payoffs are sensitive to changes in the discount factor δ.

and shirks. Suppose that Agent B, regardless of the state of the world, plays *work* with probability $\frac{3}{4}$ and *shirk* with probability $\frac{1}{4}$. This yields the following expected utility to each agent:

$$EU_A^{col} = r\left[\left(\frac{3}{4}p_1 + \frac{1}{4}p_0\right)w_A - c_A^1\right] + (1-r)\left[\left(\frac{3}{4}p_1 + \frac{1}{4}p_0\right)w_A - c_A^2\right] = 31.5 \tag{55}$$

$$EU_B^{col} = r\left[\frac{3}{4}(p_1 w_B - e) + \frac{1}{4}p_0 w_B - c_B^1\right] + (1-r)\left[\frac{3}{4}(p_1 w_B - e) + \frac{1}{4}p_0 w_B - c_B^2\right] = 4.96 \tag{56}$$

Since $EU_A^{col} > EU_A^{del}$ and $EU_B^{col} > EU_B^{del}$, both players are strictly better off under these sets of strategies compared to the principal's desired equilibrium. It remains to show that these strategies form an equilibrium of the dynamic game, given the threat of punishment (γ_1 *and shirk*; *shirk*) being played in every period and in either state.[11]

We first consider Agent A's incentive to deviate. In state ω_1, by the proof of Proposition 3, Agent A's most profitable deviation is γ_1 *and shirk*, which is his prescribed action; hence, Agent A has no short term incentive to deviate from this strategy in state ω_1. In state ω_2, if Agent B shirks, by the proof of Proposition 3 Agent A's most profitable deviation is γ_1 *and shirk*; if Agent B works, Agent A's most profitable deviation is either γ_1 *and shirk* or γ_2 *and shirk*, the latter of which is his prescribed action.[12] Altogether, for Agent A we require the following constraints to hold:

$$(1-\delta)(p_0 w_A - c_A^2) + \delta EU_A^{col} \geq (1-\delta)\left(q_0 w_A - c_A^1\right) + \delta\left[r p_0 w_A + (1-r)\, q_0 w_A - c_A^1\right] \tag{57}$$

$$(1-\delta)(p_1 w_A - c_A^2) + \delta EU_A^{col} \geq (1-\delta)\left(q_1 w_A - c_A^1\right) + \delta\left[r p_0 w_A + (1-r)\, q_0 w_A - c_A^1\right] \tag{58}$$

Note that since $p_0 - q_0 < p_1 - q_1$, if (57) holds then so does (58). Numerically, (57) reduces to $26.08 \geq 26.07$, which is satisfied.

Next, we consider Agent B's incentive to deviate. In either state, by the proof of Proposition 3, Agent B's most profitable short-term deviation is to *shirk*. Thus, we must check that Agent B would not shirk (in either state) when his prescribed action is to work; we therefore require:

$$(1-\delta)(p_1 w_B - e - c_B^1) + \delta EU_B^{coll} \geq (1-\delta)\left(p_0 w_B - c_B^1\right) + \delta\left[r p_0 w_B + (1-r)\, q_0 w_B - c_B^1\right] \tag{59}$$

$$(1-\delta)(p_1 w_B - e - c_B^2) + \delta EU_B^{coll} \geq (1-\delta)\left(p_0 w_B - c_B^2\right) + \delta\left[r p_0 w_B + (1-r)\, q_0 w_B - c_B^1\right] \tag{60}$$

Note that these constraints are equivalent. Numerically, they both reduce to $4.84 \geq 4.83$, which is also satisfied. Hence, neither player has an incentive to deviate and thus the collusive strategy outlined forms an equilibrium in which both players are strictly better off than the principal's desired equilibrium.[13]

[11] By the proof of Proposition 3, this is the unique pure strategy Nash Equilibrium of both underlying stage games.

[12] By the proof of Proposition 3, Agent A's most profitable deviation in the short term is always to shirk.

[13] Using the notation of the Proof of Proposition B1, note that we have $\Delta_{sum} = 21.05 < \widetilde{\Delta}_{sum}^1 = 31.8$. Accordingly, both players shirking yields a strictly higher surplus than both players working; this is what opens the door to collusion in this example.

Proof of Proposition B1. We initially restrict attention to cases where $\Delta_A \leq \Delta_A^{crit}$ and $\Delta_{sum} = \Delta_A + \Delta_B \geq 0$; the alternative cases are discussed at the end of the proof. We begin by deriving conditions under which the principal's desired outcome, whereby both agents work and the correct production method is implemented in each period, maximises the agents' surplus amongst all possible outcomes in both states, regardless of whether they can be implemented as an equilibrium. For a given wage scheme $\{w_A, w_B\}$, it is straightforward to verify that this is the case *iff.* the following set of constraints are satisfied:

$$(p_2 - p_1)(w_A + w_B) \geq e \tag{61}$$

$$(p_2 - p_0)(w_A + w_B) \geq 2e \tag{62}$$

$$(p_2 - q_2)(w_A + w_B) \geq -\Delta_{sum} \tag{63}$$

$$(p_2 - q_1)(w_A + w_B) \geq e - \Delta_{sum} \tag{64}$$

$$(p_2 - q_0)(w_A + w_B) \geq 2e - \Delta_{sum} \tag{65}$$

$$(p_2 - q_2)(w_A + w_B) \geq \Delta_{sum} \tag{66}$$

$$(p_2 - q_1)(w_A + w_B) \geq e + \Delta_{sum} \tag{67}$$

$$(p_2 - q_0)(w_A + w_B) \geq 2e + \Delta_{sum} \tag{68}$$

Claim 1. Let $\Delta_{sum} \geq 0$. If (62), (66) and (68) hold, all eight constraints (61)-(68) are satisfied.

Proof. If (62) holds, then:

$$(p_2 - p_1)(w_A + w_B) + (p_1 - p_0)(w_A + w_B) \geq 2e \tag{69}$$

which implies that (61) is satisfied since $p_2 - p_1 > p_1 - p_0$. Next, summing (66) and (68) and rearranging yields:

$$(2p_2 - (q_2 + q_0))(w_A + w_B) \geq 2e + 2\Delta_{sum} \tag{70}$$

Since $q_2 - q_1 > q_1 - q_0 \iff q_2 + q_0 > 2q_1$, we therefore have:

$$(2p_2 - 2q_1)(w_A + w_B) \geq 2e + 2\Delta_{sum} \tag{71}$$

which implies (67). Finally, note that since we are restricting attention to the case where $\Delta_{sum} \geq 0$, (66), (67) and (68) immediately imply (63), (64) and (65), respectively. □

Since we are restricting attention to cases where $\Delta_A \leq \Delta_A^{crit}$, using w_A^{del} and w_B^{del}, conditions (62), (66) and (68) can be rewritten as:

$$\Delta_{sum} \geq \frac{2e\left[p_0 - (1-\delta)p_1 - \delta\left[rp_0 + (1-r)\,q_0\right]\right]}{(p_2 - p_0)\,\delta\,(1-r)} =: \widetilde{\Delta}_{sum}^1 \tag{72}$$

$$\Delta_{sum} \leq \frac{(p_2 - q_2)\,2e}{p_2 - (1-\delta)p_1 - \delta\left[rp_0 + (1-r)\,q_0\right] - (p_2 - q_2)\,\delta\,(1-r)} =: \widetilde{\Delta}_{sum}^2 \tag{73}$$

$$\Delta_{sum} \leq \frac{2e\left[(1-\delta)p_1 + \delta\left[rp_0 + (1-r)\,q_0\right] - q_0\right]}{p_2 - p_2\delta\,(1-r) - (1-\delta)p_1 - \delta rp_0} =: \widetilde{\Delta}_{sum}^3 \tag{74}$$

respectively. Moreover, the principal will choose to implement delegation *iff.* the following condition holds:

$$\frac{2e + \delta\,(1-r)\,\Delta_{sum}}{p_2 - (1-\delta)p_1 - \delta\,[rp_0 + (1-r)\,q_0]} < \frac{2e}{p_2 - (1-\delta)p_1 - \delta p_0} \tag{75}$$

$$\iff \Delta_{sum} < \frac{2e\,(p_0 - q_0)}{p_2 - (1-\delta)p_1 - \delta p_0} =: \Delta^*_{sum} \tag{76}$$

That is, if total wages under delegation are strictly less than total wages under centralisation. We have the following:

$$\Delta^*_{sum} > \widetilde{\Delta}^1_{sum} \iff (1-\delta)\,(p_1 - p_0)\,[p_2 - (1-\delta)p_1 - \delta\,[rp_0 - (1-r)\,q_0]] > 0 \tag{77}$$

$$\widetilde{\Delta}^2_{sum} > \Delta^*_{sum} \iff [(p_2 - p_0) - (q_2 - q_0)]\,[p_2 - (1-\delta)p_1 - \delta\,[rp_0 + (1-r)\,q_0]] > 0 \tag{78}$$

$$\widetilde{\Delta}^3_{sum} > \Delta^*_{sum} \iff (1-\delta)\,[p_1 - p_0]\,[p_2 - (1-\delta)p_1 - \delta\,[rp_0 + (1-r)\,q_0]] > 0 \tag{79}$$

all three of which are always satisfied. In addition, we have $\Delta^*_{sum} > 0$.

Altogether, we have the following. One can always choose a combination of $\Delta_A \in \left(\widehat{\Delta}_A, \Delta^{crit}_A\right]$ and $\Delta_B < \Delta^{max}_B$ which yield a particular value of $\Delta_{sum} \in \left[0, \Delta^{crit}_A + \Delta^{max}_B\right)$. Moreover, it is straightforward to verify that $\Delta^{crit}_A + \Delta^{max}_B > \Delta^*_{sum}$; accordingly, the intersection between the intervals $\left[0, \Delta^{crit}_A + \Delta^{max}_B\right)$ and $\left[\widetilde{\Delta}^1_{sum}, \Delta^*_{sum}\right)$ is non-empty.

Select any combination of $\Delta_A \in \left(\widehat{\Delta}_A, \Delta^{crit}_A\right]$ and $\Delta_B < \Delta^{max}_B$ such that the resulting Δ_{sum} is in this intersection of intervals. By Proposition 3, delegation which induces the alternative punishment strategy to centralisation is implementable and since we have $\Delta_{sum} < \Delta^*_{sum}$, the total wage costs are strictly lower than those under centralisation. Moreover, since $\Delta_{sum} \geq \widetilde{\Delta}^1_{sum}$ and $\Delta_{sum} < \Delta^*_{sum} < \widetilde{\Delta}^2_{sum}, \widetilde{\Delta}^3_{sum}$, by Claim 1 the principal's desired outcome maximises the agents' surplus.

A formal examination of the cases where $\Delta_A > \Delta^{crit}_A$, or where $\Delta_{sum} < 0$ is not necessary for the proof. Nonetheless, one can show that in all situations where $\Delta_A > \Delta^{crit}_A$, $\Delta_{sum} > 0$ and delegation is optimal as an organisational structure, Δ_A and Δ_B can never become so high that colluding to implement production method γ_1 in both states of the world increases the joint surplus. However, the proof is significantly lengthier in this case, and is therefore omitted. One implication of this result is that, by continuity, there will always exist a non-empty set of (Δ_A, Δ_B) pairs, with $\Delta_A > \Delta^{crit}_A$, such that delegation is both optimal and collusion proof. □

References

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.

Baker, G., Gibbons, R. and Murphy, K. J. (2002), 'Relational contracts and the theory of the firm', *The Quarterly Journal of Economics* **117**(1), 39–84.

Che, Y.-K. and Yoo, S.-W. (2001), 'Optimal incentives for teams', *American Economic Review* **91**(3), 525–541.

Fudenberg, D. and Maskin, E. (1986), 'The folk theorem in repeated games with discounting or with incomplete information', *Econometrica* **54**(3), 533–554.

Glover, J. and Kim, E. (2023), On the value of demographic diversity in teams. Working Paper.

Itoh, H. (1992), 'Cooperation in hierarchical organizations: An incentive perspective', *Journal of Law, Economics, & Organization* **8**(2), 321–345.

Itoh, H. (1993), 'Coalitions, incentives, and risk sharing', *Journal of Economic Theory* **60**(2), 410–427.

Kvaløy, O. and Olsen, T. E. (2006), 'Team incentives in relational employment contracts', *Journal of Labor Economics* **24**(1), 139–169.

Levin, J. (2003), 'Relational incentive contracts', *American Economic Review* **93**(3), 835–857.