
Volume 23, issues 3–4, December 2011
ISSN 0167-6245
ELSEVIER
INFORMATION ECONOMICS AND POLICY
Available online at www.sciencedirect.com
SciVerse ScienceDirect

Information Economics and Policy 23 (2011) 252–269

Contents lists available at SciVerse ScienceDirect

Information Economics and Policy

journal homepage: www.elsevier.com/locate/iep

Simultaneous versus sequential knowledge transfer in an organization

Anthony M. Marino

Marshall School of Business, University of Southern California, Los Angeles, CA 90089-0804, United States

A R T I C L E   I N F O

Article history:
Received 26 May 2011
Accepted 29 July 2011
Available online 22 August 2011

JEL classification:
L5

Keyword:
Modes of knowledge transfer

A B S T R A C T

This paper studies an organizational knowledge sharing process which requires costly "teaching" and "learning" efforts on the part of the sender and receiver, respectively. The process is a team problem in which the principal rewards successful sharing by optimally rewarding performance. In this setting we compare two modes of knowledge transfer with regard to efficiency. The first is sequential in which the sender precommits to teaching and the receiver acts as a follower. The second is simultaneous where each agent simultaneously exerts effort. A key result is that the sequential mode dominates when teaching and learning are complements, but the simultaneous mode dominates if teaching and learning are substitutes.

1. Introduction

Today we have a wide and expanding variety of choices for modes of communication. Depending on the content of the message, we can, for example, utilize face to face communication, voice communication, E-mail, text, voice mail, a web page presentation with possible downloads and links, video chat, written chat, and hard copy written communication, just to name a few. At a general level, each of these modes can be roughly characterized as being sequential or simultaneous in nature. By sequential we mean that the sender of the information acts as a Stackelberg leader and constructs and precommits to a message intended for a receiver at a later time. Examples include E-mail and web based presentations. By simultaneous we mean that the sender sends and the receiver receives the message at the same time, as in a Nash type simultaneous move game. Related examples include face to face communication and voice communication.

This paper focuses on the mode of knowledge transfer by exploring some of the key differences between sequential and simultaneous communication in cases where knowledge sharing requires costly effort on the part of the sender and the receiver and where the idea or the process being communicated is fairly complex and not necessarily characterizable by a simple model which states that the sender has knowledge of a scalar say θ which is unknown to the receiver and which can be communicated without cost. It might be an idea and an implementation method for a cash flow production process which requires careful presentation by the sender (teacher) and serious consideration and study by the receiver (student). The teacher exerts effort to prepare presentation materials and the student would have to exert effort to completely comprehend the material. Further, in each case, the teaching and learning process could be implemented through a sequential process or through a simultaneous procedure. That is, the teacher could utilize the sequential approach by presenting the relevant materials for learning, for example, on a web site which is accessible by the student at a time of the student's choosing. This mode involves a teacher pre-committing to a level of teaching effort and a student observing or inferring that effort from the presentation before exerting learning effort. Alternatively, the teacher could adopt the simultaneous approach and conduct a face to face session with the student where the materials would be presented to and assimilated by the student (teaching and learning efforts simultaneously exerted), and there is no precommitment of teaching effort. In a model which

E-mail address: amarino@usc.edu

0167-6245/$ - see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.infoecopol.2011.07.003

highlights some of the key differences between these two approaches to teaching and learning, we want to examine the conditions under which the sequential or the simultaneous method would be more efficient. In either mode of knowledge transfer, there is a team problem in that the communication process is privately costly to both the teacher and the student, but the sending and assimilation of the information is a joint production process.[1]

One of the core tasks of any organization is to develop efficient mechanisms for knowledge sharing and knowledge transfer. There are many examples. In manufacturing, the Toyota Production System specifies protocols for knowledge transfer which make use of the simultaneous mode of teaching and learning.[2] A good example is Toyota's implementation of the kaizen process (continuous improvement and change for the better) which usually involves a small worker group attempting to improve their own work environment and productivity with leadership being provided by a line supervisor. My academic department has instituted a mentoring system for junior faculty which employs the simultaneous approach. A senior faculty member meets directly with a junior faculty to convey information on teaching methods and materials and the research process. It is well known that software development firms make extensive use of the sequential mode of knowledge transfer. In consulting, Bain and Company has made use of the sequential mode since the 1980s by creating a database of summaries of consulting projects. Xerox's "Eureka" system is a sequential mode repair database for knowledge sharing that saves the company millions a year in repair costs.[3] In each of the above examples, employees were incentivized to participate in the knowledge sharing process through monetary or in-kind compensation.

Our basic model focuses on such organizational knowledge transfer. We characterize an organization consisting of a principal and two agents. One agent may be informed about a fairly complicated idea which is capable of generating additional cash flow for the firm, and the other agent may not be informed. The principal develops an optimal incentive scheme to encourage the informed agent to attempt to educate the uninformed agent about the new process. The attempt at this knowledge transfer requires costly effort both on the part of the sender (teacher) and the receiver (student or learner). We characterize this process at the principal's optimal contracts under the sequential and simultaneous modes of communication. Next we compare the principal's expected equilibrium profit under the two alternative communication methods. We show that if teaching and learning efforts are complements in the communication process, then the sequential mode dominates, but if these efforts are substitutes, the simultaneous mode dominates. These results hold true in a model where the basic technology of knowledge sharing is the same across modes and only the sequentiality or simultaneity of one process versus the other is isolated and characterized.

Next, we extend these results by considering other key differences between the two approaches. The asynchronous nature of the sequential mode enables the receiver to access information at a time and place of his/her choice. This allows the receiver to attempt to assimilate information when effort cost is low and, thus, efficiency is increased. The simultaneous mode does not share this advantage, but it does have the advantage of allowing the teacher to locally adapt to the ability and the effort cost of the student at hand. The sequential mode does not permit such local adaptation. These two features are added to the basic model and the relative efficiency of the two approaches is studied. A final section of the paper presents two extensions.

The economics literature on communication has to some extent ignored the agent-to-agent presentation and assimilation of information as an explicit production process. The paper by Dewatripont and Tirole (2005) is the closest to this paper. They present a model of costly communication in which both the sender and receiver of information exert costly effort to send and assimilate, respectively, a piece of knowledge. They model imperfect congruence between the sender and receiver and formulate the problem as one involving moral hazard in teams. Their paper then characterizes communication equilibria which depend on the level of congruence between the sender and receiver, the nature of decision making, and the knowledge that the sender has about the receiver's payoffs. The key results employ simultaneous move games where sending and receiving efforts are strategic complements. In an extensions section, they briefly discuss the case of strategic substitutes and how congruence affects total effort in equilibrium. Also, in this section they touch on the effects of sequential communication. In contrast, our paper assumes that there is complete congruence through a principal's optimal contract, it fixes the nature of decision making and it assumes symmetric knowledge of payoffs. Our focus is only on the mode of knowledge transfer in the context of the team problem and its effects on the second best equilibrium at the principal's optimal contract. The survey paper by Van Zandt (1999) nicely summarizes the large literature on information processing and dissemination within firms. This literature looks at firm level processing as opposed to communication between agents with incentive issues. Dessein and Santos (2006) endogenize the firm's choice of how much to let agents make use of local information (adaptiveness). That is, they endogenize the quality of information in the firm. In our paper, the quality of the communication process is decided by the communicating individuals, given an incentive contract designed by the principal. Finally, a related paper by Itoh (1991) studies the incentives of agents to help other agents. The principal selects an optimal compensation scheme which results in a task structure whereby agents either specialize effort in their own tasks or are motivated to help other agents. He provides a sufficient condition for helping to be optimal. Our communication structure is analogous to Itoh's notion of helping, but instead of focusing on the incentives of agents to engage in helping or not, we concentrate on how to help. That is, we are interested in studying the mode of helping between agents.

1 See Marschak and Radner (1972).

2 See Dryer and Nobeoka (2000).

3 See Thurm (2006).

Section 2 presents the basic model and the two alternative approaches to knowledge transfer. Section 3 compares the two modes of communication with respect to the principal's expected profit. Section 4 introduces features other than sequentiality and simultaneity distinguishing the two modes and conducts a comparison. Section 5 discusses two extensions. The first considers the case where the firm markets its knowledge to an external market by using the sequential approach, and the second endogenizes local adaptation in the simultaneous approach. Section 6 concludes.

2. The basic model

2.1. The two division firm and costly knowledge sharing

An organization has two divisions each of which has a division manager. The divisions and their managers are identical. The principal contracts with both managers. A manager controls a cash flow process which is subject to randomness. With probability π (independent across managers), a manager i is endowed with knowledge of a cash production function given by

$$y_i = f(x_i) - x_i, \quad i = 1,2. \tag{1}$$

The variable x_i is a scale variable chosen by the manager and the function f is assumed to be strictly concave and satisfy

$$f' > 0, \quad f'' < 0, \quad f'(0) = \infty, \quad f'(\infty) = 0, \quad \text{and } f(0) = 0.$$

The manager's choice of x_i, denoted x^*, is given by the solution to

$$\max_{\{x_i\}} f(x_i) - x_i,$$

or $f'(x^*) - 1 = 0$. Define y as $y = f(x^*) - x^*$. It is assumed that a manager with such information can costlessly implement x^* and y.[4] With probability $(1 - \pi)$ a manager receives no information and can produce nothing. Let I denote the situation that either manager is informed of 1 and let U denote the situation that a manager is uninformed. There are then four information states say (i,j) for the two managers, where i represents manager 1's state of informativeness and j represents manager 2's state of informativeness. The two states (I,I) and (U,U) are those in which either both managers are endowed with the information to produce a positive cash flow or both managers are uninformed, respectively. These occur with probabilities π^2 and $(1 - \pi)^2$, respectively, and require no communication.[5] In the other two states, (I,U) and (U,I), one manager is informed and the other is uninformed, and these occur with probability $(1 - \pi)\pi$. In these two states, it can be profitable to use costly communication to transmit information from an informed manager to an uninformed one. Managers observe the state of informativeness of the other manager, but the principal does not observe this information.

The process whereby one manager communicates the cash flow production process to the other is described by the function $p(T_i, L_j)$, $p : D_s \to [0,1]$, $D_s \subset \mathbb{R}^2_+$, which measures the probability that agent i will perfectly teach or communicate the knowledge that he received as described by (1) to the other agent. The variable T_i denotes the teaching effort of the informed agent and L_j denotes the learning effort of the uninformed agent. Let subscripts denote partial derivatives. We assume throughout that p satisfies the following assumptions.

A.0 p is twice continuously differentiable in both variables with $p_{T_i}, \ p_{L_j} > 0$, and p has a negative definite Hessian at each point in its domain. Further, $p_{T_iL_j}$ is globally positive or globally negative. In addition, we make use of the more specific quadratic functional form.

A.1 $p(T_i, L_j) \equiv T_i + L_j - T_i^2 - L_j^2 + sT_iL_j$, where $s \in \{-1, +1\}$. Given $s \in \{-1, +1\}$, $D_s = \{(T_i, L_j) | p \in [0, 1)$ and $p_{T_i}, \ p_{L_j} > 0\}$.

The quadratic function p is strictly concave and increasing in T_i and L_j in its domain D_s with an interaction term T_iL_j. If the interaction term is positive, T_i and L_j are complements and they are substitutes if it is negative. That is, teaching and learning are either global complements or global substitutes. More teaching effort might increase the marginal productivity of learning effort through the teacher clarifying and framing issues in a more compelling manner so as to make extra learning effort more productive ($\partial^2 p(T_i, L_j)/\partial L_j \partial T_i > 0$). Alternatively, more teaching effort could crowd out learning effort because receivers of information can learn more on the margin through a more learning effort intensive mix. The student must do relatively more for herself or himself in terms of effort in order to increase the marginal benefit of learning ($\partial^2 p(T_i, L_j)/\partial L_j \partial T_i < 0$). A key feature of this model which makes it different from most rivalrous games is that increased teaching or learning effort by one player each increase the payoff to the player exerting effort and to the other player. However, while teaching and learning each increase total payoff, they can be strategic complements or substitutes when one examines the effect of an increase in one of these variables on the marginal payoff of the other.[6] On the cost side, we assume that each unit of teaching or learning effort exerted by a manager carries with it a one dollar effort cost.

The sequence of decisions begins at stage 0 where the principal hires and contracts with the agents. All agents are risk neutral and the principal cannot observe or contract on teaching or learning efforts nor can the principal contract on the act of teaching or learning. The principal contracts with the agents based only on observed returns. Suppose that the optimal contract necessitates that the principal pays a positive contingent payment only if both divisions produce the returns y. At an intermediate stage

[4] The key point is that knowledge produces payoff $y > 0$. This example is presented to emphasize that the knowledge being transferred represents a complex process.

[5] An informed manager of either type is informed with the same basic knowledge as given by (1), so that there is no need to share knowledge if the state is (I,I).

[6] See Bulow et al. (1985) for a discussion of strategic complements and substitutes in rivalrous market games.

1, agents receive or do not receive information regarding the process (1). Given the principal's payment scheme, agents in the states (I,U) and (U,I) are motivated to exert effort in teaching and learning. At stage 2 after all possible teaching and learning take place, uninformed managers do nothing, informed managers implement scales x^* and generate returns y, and payments by the principal are made. We assume that if an informed manager exerts effort to communicate information to the other manager and fails, then that manager will not be paid but will costlessly generate y, as he is indifferent between this action and doing nothing. Under this assumption, it is clear that it is optimal for the principal to make a payment to an agent only if both agents generate positive returns.

We will consider two modes of knowledge sharing with respect to the technology p. The first is a simultaneous move technology in which the teaching and learning agents come face to face and simultaneously exert their efforts to determine a Nash equilibrium in efforts. We think of this as an interactive technology and it is meant to describe situations where an informed manager explains the production process (1) in a one-on-one manner to the uninformed manager.

The second technology is sequential in nature in that the informed agent acts first as a Stackelberg leader and exerts teaching effort. The uninformed agent observes teaching effort and then follows its reaction function in exerting learning effort. This mode of knowledge sharing is meant to capture situations where the manager with information attempts to explain or communicate that information to the other division managers by preparing a memo or E-mail, drawing up a page in a manual, or posting an explanation of the information on a board or on the web. An uninformed manager can then attempt to learn that information by reading and understanding it from the posting place. A crucial assumption is that the potential learner can observe or infer teaching effort and this seems reasonable for the key examples that the model is designed to illustrate. For example, in a web based presentation or even in a hard-copy programed learning guide used in distance learning, the learner can discern or infer the effort put in by the designer when they execute the lesson plan.

We will formulate these modes using the same technology p, in an effort to study the pure effects of simultaneous as opposed to sequential knowledge transfer. Later we will introduce some appropriate asymmetries in the technologies of communication which can appear when comparing sequential versus simultaneous modes.

2.2. A simultaneous move or interactive technology

The principal pays a division manager a salary S and a contingent compensation c if both divisions achieve returns y. Agents follow Nash strategies. The utility or net expected compensation of the manager i is written as W_i and given by

$$W_i = S + c\{\pi^2 + \pi(1-\pi)[p(T_i, L_j) + p(T_j, L_i)]\} - \pi(1-\pi)(T_i + L_i).$$

An informed agent i will choose his or her teaching effort as the solution to

$$c[\partial p(T_i, L_j)/\partial T_i] - 1 = 0, \tag{2}$$

and an uninformed agent j who is being informed will choose learning effort as the solution to

$$c[\partial p(T_i, L_j)/\partial L_j] - 1 = 0. \tag{3}$$

Thus, given c and the other agent's choice of effort, each agent sets the marginal benefit of teaching or learning equal to its marginal cost which is a dollar. For convenience in what follows, we take the manager's reservation utility to be zero, so that the relevant participation constraint is

$$W_i \geqslant 0. \tag{4}$$

Finally, we assume that the manager has limited wealth so that the limited liability constraints

$$c \geqslant 0, \quad S \geqslant 0 \tag{5}$$

hold.

The principal will want to design an incentive contract subject to the incentive compatibility, (2) and (3), participation, (4), and limited liability constraints, (5), so as to maximize the objective function

$$2\pi^2(y-c) + 2(1-\pi)\pi p(T_1, L_2)(y-c) + 2(1-\pi)\pi p(T_2, L_1)(y-c) + (1-\pi)\pi[1 - p(T_1, L_2)]y + (1-\pi)\pi[1 - p(T_2, L_1)]y - 2S.$$

Note that if an agent is informed and fails to communicate successfully, he alone generates y and the principal does not pay c. This state occurs with probability $1 - p(T_i,L_j)$, conditional on one agent being informed and the other agent not being informed. The principal's objective function simplifies to

$$2(1-\pi)\pi y + 2\pi^2(y-c) + (1-\pi)\pi p(T_1, L_2)(y-2c) + (1-\pi)\pi p(T_2, L_1)(y-2c) - 2S. \tag{6}$$

We assume that the participation constraints are nonbinding and that the incentive compatibility constraints are binding at an interior solution. We will introduce sufficiency conditions later which guarantee that such an interior solution exists. The appendix formulates the principal's problem and shows that under these assumptions, the optimal non-contingent payment, S, is zero.

Define the $T_i^n(c)$ and $L_j^n(c)$ as the solutions to the incentive compatibility constraints, (2) and (3). Next, define the function $p^n(c) = p\left(T_i^n(c), L_j^n(c)\right)$ corresponding to the two identical teaching and learning agents. This reduced form function can be used to further characterize the principal's optimal selection of c. With the participation constraints nonbinding, the principal's problem is characterized by

$$\max_{\{c\}} \Omega^n(c), \quad \text{where } \Omega^n(c) = 2(1-\pi)\pi y + 2\pi^2(y-c) + 2(1-\pi)\pi p^n(c)(y-2c). \tag{7}$$

In the case of our quadratic formulation (1), the incentive compatibility constraints generate

$$T_i^n(c),\ L_j^n(c) = \frac{(c-1)}{c(2-s)}, \quad \text{and} \quad p^n(c) = \frac{1}{c^2}\frac{(c^2-1)}{(2-s)},$$
$$s \in \{-1, 1\}.$$

We have (all proofs are provided in the appendix.)

Lemma 1. *Let A.1 hold and $y > \frac{4-\pi}{2(1-\pi)}$, then the reduced form problem (7) has a unique interior solution. At such a solution the principal's optimal payment satisfies $1 < c < y/2$ and there is positive expected profit.*

If the optimal reward for communication satisfies $c < y/2$, then the incentive compatibility constraints for T_i and L_j imply that

$$y\left[\partial p T_i^n(c), L_j^n(c)/\partial T_i\right] - 1 > 0$$
$$\text{and } y\left[\partial p T_i^n(c), L_j^n(c)/\partial L_j\right] - 1 > 0. \tag{8}$$

At a first best benchmark, the left sides of Eq. (8) are zero, so that we have shown that, at the second best contract, teaching and learning efforts are under supplied.

The results for the simultaneous mode of communication are summarized in Proposition 1.

Proposition 1. *Let A.1 be satisfied and let the assumptions of Lemma 1 hold. The second best contract under a simultaneous communication scheme pays a positive wage to a division manager only if all managers generate the highest return. At the optimal compensation scheme, both teaching and learning efforts are under supplied by the manager. Further, increases in π generate a smaller optimal c and greater expected profit for the principal.*

The only state of the world in which the principal must encourage the informed and uninformed agents to exert effort is the state in which one agent is informed (the potential teacher) and the other is not informed (the potential student). In this state, the principal optimally spends a total of $2c$ to get both agents to exert effort, and the principal receives y on the margin in return for this expenditure. In this state, there is a team problem, because it takes two agents exerting effort to make possible a single extra y. Given that the limited liability constraints are binding, it is not profitable for the principal to raise c to the point where $c = y$ and the first best efforts are obtained. From (6), the part of the principal's objective function impacted by variations in c is $2\pi^2(y-c) + 2(1-\pi)\ \pi p^n(c)(y-2c)$ and at $c = y$ this is given by $-2\pi p^n(y)\ y(1-\pi) < 0$.

Finally, at the second best, a greater exogenous probability that an agent will be informed will raise expected profit to the principal and, at the same time, lower the optimal payment to incentivize communication. That is, organizations employing more creative and educated individuals (assuming that these types are more prone to originating new ideas) can incentivize simultaneous communication at a lower cost than those not employing these types. New ideas are optimally spread at a lower cost if the organization has more talent. What drives this compensation result is that the principal pays c in state (I,I), occurring with probability π^2, and states (I,U), (U,I), occurring with probability $(1-\pi)\pi$. In the former state more c does not generate extra cash flow but in the latter states it does. If π increases, π^2 strictly increases and this induces the principal to reduce c, while $\pi(1-\pi)$ increases for $\pi < 0.5$ and decreases otherwise. The principal's c reductions in state (I,I) swamps any incentive to raise c for incentive reasons in states (I,U), (U,I).

2.3. A sequential move or posting technology

In this section, we augment the above model by changing the mode of information sharing. A manager with information can attempt to explain or communicate that information to other division managers by preparing a memo, drawing up a page in a manual, or posting an explanation of the information on a board or on the web. These are all mechanisms of precommitment. An uninformed manager can then attempt to learn that information by reading and understanding it from the posting place. This teaching and learning requires costly effort on both the posting and retrieving sides of the transfer. The new dimension in this version of the problem is that a poster of information is a Stackelberg leader and a retriever of information is a follower. The display of the posted information by the teacher allows the learner–follower to infer teaching effort. Thus, each agent is a leader when posting and a follower when learning. So that the two processes are directly comparable, let us characterize the posting and learning technology using the same function $p(T_i, L_j)$.

An uninformed agent j chooses learning effort as a follower when learning information that has been posted by agent i.

$$\max_{\{L_j\}}\ cp(T_i, L_j) - L_j.$$

The solution to this problem entails

$$c\partial p(T_i, L_j)/\partial L_j - 1 = 0. \tag{9}$$

which in turn generates the reaction function $L_j(T_i, c)$ with

$$\partial L_j(T_i,c)/\partial T_i = \frac{\partial^2 p(T_i,L_j)/\partial L_j \partial T_i}{-[\partial^2 p(T_i,L_j)/\partial^2 L_j]} \text{ and } \partial L_j(T_i,c)/\partial c$$
$$= \frac{\partial p(T_i,L_j)/\partial L_j}{-[\partial^2 p(T_i,L_j)/\partial^2 L_j]} > 0.$$

The informed agent solves

$$\max_{\{T_i\}} cp(T_i, L_j(T_i, c)) - T_i,$$

and the solution entails

$$c[\partial p(T_i, L_j(T_i,c))/\partial T_i$$
$$+ (\partial L_j(T_i,c)/\partial T_i)(\partial p(T_i, L_j(T_i,c))/\partial L_j)] - 1 = 0. \tag{10}$$

If we assume that the participation constraints are non-binding, that the incentive compatibility constraints are binding, and that there is an interior solution (we will give sufficiency conditions later), then we show in the appendix that $S = 0$. As in the simultaneous case, we can write the principal's problem in reduced form and study existence and uniqueness of the solution. The solution to (10) gives us $T_i^s(c)$ which can be substituted into (9) to yield $L_j(T_i(c), c) = L_j^s(c)$. Next define the reduced form function

$$p^s(c) = p\left(T_i^s(c), L_j^s(c)\right). \tag{11}$$

With the participation constraints nonbinding, the principal's reduced form problem is characterized by

$$\max_{\{c\}} \Omega^s(c), \text{ where } \Omega^s(c) = 2(1-\pi)\pi y + 2\pi^2(y-c) + 2(1-\pi)\pi p^s(c)(y-2c). \tag{12}$$

The objective function (12) is identical in structure to that of (7) except for the sequential probability function $p^s(c)$ replacing the simultaneous function $p^n(c)$. For our quadratic case,

$$T_i^s(c) = \frac{1}{3c}(2(c-1)+cs), \quad L_j^s(c) = \frac{1}{6c}(4c-2s+2cs-3), \quad \text{and } p^s(c) = \frac{1}{12c^2}(4c^2s+8c^2+1-8).$$

Results similar to those of Lemma 1 hold true in the present context. We have

Lemma 2. *Let A.1 hold and $y > \frac{19-13\pi}{7(1-\pi)}$, then the reduced form problem (12) has a unique interior solution. At such a solution the principal's optimal payment satisfies $1 < c < y/2$. and there is positive expected profit.*

We can use the bounds provided by Lemma 2 to determine whether, at the optimal reward for communication, teaching and learning efforts are under or over supplied. For learning effort, the incentive compatibility condition (9) and the fact that $c < y/2$ imply that

$$y\partial p\left(T_i^s(c), L_j^s(c)\right)/\partial L_j > 1, \tag{13}$$

so that learning effort is under supplied at optimum. In the case of substitutes, it is clear from the incentive compatibility condition (10) that

$$y\partial p\left(T_i^s(c), L_j^s(c)\right)/\partial T_i > 1, \tag{14}$$

by $c < y/2$ and $(\partial L_j(T_i,c)/\partial T_i)(\partial p(T_i, L_j(T_i,c))/\partial L_j) < 0$. In the case of complements, the latter term is positive and to see that, in fact, (14) holds note that

$$y\partial p\left(T_i^s(c), L_j^s(c)\right)/\partial T_i = y\left(1 - 2\left(\frac{1}{3c}(3c-2)\right) + \frac{1}{6c}(6c-5)\right) = \frac{1}{2c}y > 1. \tag{15}$$

In the complements case there is under supply of teaching and learning efforts at the sequential scheme given the principal's optimal compensation. We can summarize the sequential move case in

Proposition 2. *Let A.1 be satisfied and let the assumptions of Lemma 2 hold. The second best contract under the sequential communication scheme pays a positive wage to a division manager only if all managers generate the highest return. At the optimal compensation scheme, both teaching and learning efforts are under supplied by the manager. Further, increases in π generate a smaller optimal c and greater expected profit for the principal.*

At the sequential mode equilibrium there is optimal under supply of both teaching and learning efforts. For the case where teaching and learning are substitutes, this result is intuitive because the learning agent follows a Nash strategy given teaching effort and the principal pays less than y in equilibrium. The teacher internalizes the marginal benefit of his own teaching effort but this is countered by the negative marginal term accounting for the decrease in learning effort that the increase in teaching effort is expected to generate. From (10), the sum of these terms is equal to 1 and this makes the basic private marginal benefit $c\partial p(T_i^s(c), L_j^s(c))/\partial T_i$ even greater than 1. Given $c < y$, the teacher under supplies effort. The case of complements is less transparent because the marginal private benefit of the teacher's own effort, $c\partial p\left(T_i^s(c), L_j^s(c)\right)/\partial T_i$, is added to a positive marginal term accounting for the increase in learning effort that the increase in teaching effort causes. From (15), the firm's basic marginal benefit $y\partial p\left(T_i^s(c), L_j^s(c)\right)/\partial T_i$ is, however, greater than one when c is optimally chosen such that $y/2 > c$. Finally, the result that a greater probability of informativeness increases the principal's profit and decreases the optimal payment holds again as in the simultaneous case. Again, more talent in the set of managers (a greater π) makes it less costly for the principal to optimally incentivize knowledge transfer. The intuition for this result is the same as in the simultaneous case.

3. A comparison of the simultaneous and sequential move technologies

In this section we compare the principal's welfare at an optimal contract under the two alternative modes of knowledge transfer. Given that the basic teaching-assimilation technology, $p(\cdot)$, is the same across modes, this analysis then isolates the pure effects of sequential versus simultaneous communication. It turns out that the principal fares better under the sequential mode if teaching and learning efforts are complements and the simultaneous mode is better if they are substitutes.

We have Proposition 3.

Proposition 3. *Let A.1 and the assumptions of Lemmas 1 and 2 hold. Suppose that teaching and learning efforts are complements (substitutes). In a second best equilibrium, the principal's expected profit is greater (less) under sequential communication than under simultaneous communication. Further, the principal's optimal compensation under the simultaneous mode of communication, c^n, is greater (less) than the principal's optimal compensation under the sequential mode of communication, c^s.*

The intuition behind the case of complements is compelling and follows from an application of strategic complements in Nash versus Stackelberg equilibria. The teacher–leader knows that another unit of teaching effort will induce more learning effort which in turn raises the marginal product of teaching. In the Nash equilibrium, the teacher does not internalize this extra boost in effort by the learner, because he assumes that the student will retain a given level of learning effort. Thus, when we compare Stackelberg and Nash teaching efforts, Stackelberg results in more teaching effort. Because the learner–follower's reaction function is strictly increasing, greater teaching effort at the Stackelberg equilibrium results in greater learning effort as compared to Nash. The principal's equilibrium expected profit is increasing in the equilibrium probability p and by generating more of both types of effort, the sequential mode delivers greater expected profit for the principal. Further because of the above effort boosting process triggered by the sequential mechanism and complementarity, the principal can implement communication under the sequential approach at a lower incentive payment.

When teaching and learning are substitutes in the knowledge transfer process, the opposite result obtains and the simultaneous communication mode generates greater profit for the principal. In this case, the learner has a downward sloping reaction function. Another unit of teaching effort by the teacher–leader then induces less learning effort and the teacher internalizes this effect, resulting in less teaching effort. A teacher using the simultaneous mode does not internalize this negative effect and teaching effort under this mode is greater. Under our assumption A.1, the direct effects of diminishing returns dominate the cross effects of diminishing marginal productivity (i.e., if $\left|\frac{p_{ii}}{p_i}\right| > \left|\frac{p_{ij}}{p_j}\right|, i \neq j$, in the region where the incentive compatibility constraints are binding). This implies that $p(\cdot)$ follows the movement of teaching effort and the sequential approach with lower teaching effort results in a lower $p(\cdot)$. The greater $p(\cdot)$ under the simultaneous approach then makes it less costly to elicit communication.

Propositions 3 holds true in more general settings than the quadratic case considered here. We have Proposition 4.

Proposition 4. *Let A.0 hold and let $p(T_i, L_j(T_i, c))$ be strictly concave in T_i, $\frac{\partial^2 p(T_i, L_j(T_i,c))}{\partial T_i^2} < 0$. Further assume that $\left|\frac{p_{ii}}{p_i}\right| > \left|\frac{p_{ij}}{p_j}\right|$, $i \neq j$. In a second best equilibrium, if teaching and learning efforts are complements (substitutes), the principal's expected profit is greater under the sequential (simultaneous) mode of communication.*

The key condition of Proposition 4 says that the direct effect of diminishing returns on own marginal productivity is greater in absolute value than the cross effect on the other marginal productivity, when an effort level is increased. For example, in the case where T_i and L_j are complements, an increase in T_i causes a percentage reduction in the marginal product of T_i which is greater in absolute value than the percentage increase in the marginal product of L_j caused by the increase in T_i. The symmetric effect is true with respect to changes in L_j. Indeed these conditions hold for our quadratic case in the feasible region.

Assume that all other factors are equal, when we compare the two methods. The above results indicate that in situations where teaching and learning are complements, we should see that the sequential method is used to communicate and where they are substitutes, the simultaneous method is used.

4. Non-symmetric comparisons of the simultaneous and sequential move technologies

So far we have assumed that all modes of communication have the same technology and effort costs, because we wanted to focus on the pure effects of the sequential versus simultaneous mechanisms. Depending on the actual situation there can be some key differences in the costs and benefits of teaching and learning efforts. We want to highlight a few that we believe are of first order importance.

One of the advantages of the sequential approach is that teaching and learning can occur asynchronously so as to permit the learner to access and assimilate the information at a time and location such that his effort cost is relatively low. We access our E-mail at the most opportune times, and we study web presentations when we would be most receptive to and productive at this learning activity. If potential students have high and low learning effort costs, depending on the state of the world, then we would expect students to access information when their effort cost is low. To model this type of self selection, which is possible under the sequential approach but not under the simultaneous approach, we assume that potential students can differ with respect to their effort cost which we denote as β_i, $i = 1, 2$, with $\beta_2 > \beta_1$. Let us normalize the low learning effort cost to that of teaching effort and set it as $\beta_1 = 1$. The probability of effort cost β_1 is denoted b.

On the other hand, a sequential presentation has a disadvantage relative to the simultaneous approach in that it is typically designed for some generic student and a generic teaching situation.[7] The teaching message is standardized and it uses none of the local information that would be used in face to face communication. The simultaneous approach does not have this problem, because the teacher's presentation can be customized to fit the local environment and student at hand. The actual mechanism might be that, when student and teacher meet, there is a possible vector of information that, if known, could be shared. Such information would clarify the communication process. Some of the elements of the vector the student knows and others he does not know and likewise for the teacher. The student can communicate to the teacher if it is not necessary for the teacher to dwell on known components of the vector, and the student can request that the teacher emphasize the unknown components of the information vector. The ability of the teacher to adapt to the needs of the student when communication is simultaneous makes this technology more efficient.[8] A reduced form model which captures this difference can

[7] A solution to this problem might be for the teacher to post a menu of presentations for all possible teaching and learning environments. However, it could be very costly for the teacher to prepare all possible presentations, or it could be costly for a learner to choose from a menu. We ignore this approach. However, no matter how extensive is a menu of presentations, it cannot replicate an "in class" presentation in terms of its

[8] On the other hand it may be that, although the simultaneous move case is more productive, it may involve more set up costs as both participants must be present when communication takes place. The sequential move technology can involve asynchronous participation. We assume that the increased efficiency effects outweigh this effect.

be formulated by placing a multiplicative local adaptation parameter $\alpha > 1$ on the simultaneous move probability function, p. The idea is that with probability $\alpha p > p$ the message will be communicated and assimilated under a simultaneous approach, whereas the sequential process has the probability function p. Relative to the sequential approach, local adaptation makes the clarity of communication greater.[9]

Let us begin with the sequential approach. Let L_{jk} denote the effort of a learning agent j with effort cost k and let T_{ik} be the corresponding teaching effort associated with this student. In the sequential case, only the low effort cost is experienced by learning agents so that learning effort is the solution to

$$c\partial p(T_{i1}, L_{j1})/\partial L_{j1} - 1 = 0.$$

A teacher solves

$$c[\partial p(T_{i1}, L_{j1}(T_{i1}, c))/\partial T_{i1} + (\partial L_{j1}(T_{i1}, c)/\partial T_{i1}) \times (\partial p(T_{i1}, L_{j1}(T_{i1}, c))/\partial L_{j1})] - 1 = 0.$$

The solution to the principal's problem is identical to that in the basic model with (T_{i1}, L_{j1}) replacing (T_i, L_j). Thus, the same reduced form functions $T^s_{i1}(c) = T^s_i(c)$ and $L^s_{j1}(c) = L^s_j(c)$ result.

In the simultaneous approach a teacher will face one of two different learning effort costs when attempting to communicate. A teacher solves

$$c[\alpha\partial p(T_{ik}, L_{jk})/\partial T_{ik}] - 1 = 0, \quad i \neq j = 1, 2, \ k = 1, 2, \tag{16}$$

and a learner solves

$$c[\alpha\partial p(T_{ik}, L_{jk})/\partial L_{jk}] - \beta_k = 0, \quad i \neq j = 1, 2, \ k = 1, 2, \ \text{where } \beta_1 = 1, \text{ and } \beta_2 > 1. \tag{17}$$

Let $\left(T^n_{ik}(c, a, \beta_k), L^n_{jk}(c, a, \beta_k)\right)$ denote the solution to (16) and (17). Routine comparative statics of the system (16) and (17) yield the following results.

Lemma 3. *Let A.0 hold and let* $\left|\frac{p_{ii}}{p_i}\right| > \left|\frac{p_{ij}}{p_j}\right|$, $i \neq j$, *be satisfied. Then we have*

(i) *If teaching and learning are complements, then* $\partial T^n_{i2}/\partial\beta_2$, $\partial L^n_{j2}/\partial\beta_2$, $\partial p\left(T^n_{i2}, L^n_{j2}\right)/\partial\beta_2 < 0$, *whereas if they are substitutes* $\partial L^n_{j2}/\partial\beta_2$, $\partial p\left(T^n_{i2}, L^n_{j2}\right)/\partial\beta_2 < 0$, *and* $\partial T^n_{i2}/\partial\beta_2 > 0$.

(ii) $\partial T^n_{ik}/\partial\alpha$, $\partial L^n_{jk}/\partial\alpha$, $\partial p\left(T^n_{ik}, L^n_{jk}\right)/\partial\alpha > 0$.

The quadratic formulation (A.1) satisfies the conditions of Lemma 3 in the region where the incentive compatibility constraints are met, so that the results hold for this case.

If we assume that the participation constraints are non-binding, that the incentive compatibility constraints are binding, and that an interior solution exists (we will give sufficiency conditions later), then we show in the appendix that $S = 0$. Define the function

$$Ep^n(c, \alpha, \beta_2) = b\alpha p\left(T^n_{i1}(c, \alpha, 1), L^n_{j1}(c, \alpha, 1)\right) + (1-b)\alpha p\left(T^n_{i2}(c, \alpha, \beta_2), L^n_{j2}(c, \alpha, \beta_2)\right)$$

which is a reduced form for the two identical teaching and learning agents taken as an expectation across the two effort cost types. The reduced form version of the principal's problem is

$$\max_{\{c\}} \Omega^n(c), \quad \text{where } \Omega^n(c) = 2(1-\pi)\pi y + 2\pi^2(y - c) + 2(1-\pi)\pi Ep^n(c, \alpha, \beta_2)(y - 2c). \tag{18}$$

Our quadratic formulation yields

$$T^n_{i1}(c, \alpha, 1), \quad L^n_{j1}(c, \alpha, 1) = \frac{c\alpha - 1}{3c\alpha}(2 + s),$$

$$\left(T^n_{i2}(c, a, \beta_2), L^n_{j2}(c, a, \beta_2)\right) = \left(\frac{c\alpha(2+s) - 2 - s\beta_2}{3c\alpha}, \frac{s(c\alpha - 1) - 2(\beta_2 - c\alpha)}{3c\alpha}\right), \text{ and}$$

$$Ep^n(c, \alpha, \beta_2) = \frac{1}{3c^2\alpha^2}\left(-b(1+s) + c^2\alpha^2(2+s) - s\beta_2(1-b) - \beta_2^2(1-b) - 1\right).$$

We have

Lemma 4. *Let A.1 hold and*

$$y > \max\left\{\left(\left(\frac{8}{3\alpha^2}\frac{1-\pi}{2-\pi}\right)\left(2b + \beta_2(1-b) + \beta_2^2(1-b) + 1\right)\right)^{1/2}, \times \left(\frac{8}{\alpha^2(2+\pi)}((1-\pi)(1 + \beta_2(\beta_2 - 1)(1+b))\right)^{1/2}\right\},$$

then the reduced form problem (18) has a unique interior solution. At such a solution the principal's optimal payment satisfies $0 < c < y/2$ and there is positive expected profit.

If the optimal reward for communication satisfies $c < y/2$, then the incentive compatibility constraints for T_i and L_j imply that

$$y[\alpha\partial p(T_{ik}, L_{jk})/\partial T_{i1}] - 1 > 0, \quad i \neq j = 1, 2, \ k = 1, 2, \tag{19}$$

$$y[\alpha\partial p(T_{ik}, L_{jk})/\partial L_{jk}] - \beta_k > 0, \quad i \neq j = 1, 2, \ k = 1, 2, \ \text{where } \beta_1 = 1, \text{ and } \beta_2 > 1. \tag{20}$$

Thus, there is under supply of efforts at optimum under our assumptions. We can extend the other results of Proposition 1 to the present simultaneous case.

Proposition 5. *Let A.1 be satisfied and let the assumptions of Lemma 4 hold. The second best contract under the revised simultaneous communication scheme pays a positive wage to a division manager only if all managers generate the highest*

[9] A real world example of this effect is pointed out by Thurm (2006). Executives at London's water supplier attempted to increase efficiency by giving inspectors hand held computers for sequential communication and eliminating the dispatching station. Efficiency soon fell as the depot was a facilitator for the face to face communication of knowledge. The need for face to face (simultaneous) communication was so great that inspectors began meeting after hours to share information at a local restaurant.

return. At the optimal compensation scheme, both teaching and learning efforts are under supplied by the manager. Further, increases in π, b, and α generate a smaller optimal c and greater expected profit for the principal, while increases in β_2 generate a greater optimal c and less expected profit for the principal.

From Proposition 3, the crucial condition determining whether from an ex ante view the sequential or the simultaneous mode of communication should be used is the following

$$p\left(T^s_{i1}(c), L^s_{j1}(c)\right) \gtreqless b\alpha p\left(T^n_{i1}(c,\alpha,1), L^n_{j1}(c,\alpha,1)\right) + (1-b)\alpha p\left(T^n_{i2}(c,\alpha,\beta_2), L^n_{j2}(c,\alpha,\beta_2)\right), \tag{21}$$

for all $c < y/2$ such that the incentive compatibility constraints hold and for $i \neq j = 1, 2$. The left side of (21) shows the equilibrium probability of successful communication under the sequential approach, at a given c, and the right side shows the expected probability of successful communication under the simultaneous approach at the same c. The latter expectation is taken over both high and low effort cost learners, because under the simultaneous approach the potential teacher does not know whether he will face a learner with high or low cost. With the sequential approach a potential teacher knows ex ante that a prospective learner will be low cost.

From the results of Lemma 4, the right side of inequality (21) is increasing in α, decreasing in β_2, and increasing in b. The simultaneous approach will dominate the sequential approach in situations where teaching and learning are substitutes, where the local adaptation parameter is large, where the likelihood that the student will have low effort cost is high or where the magnitude of the high effort cost is relatively low. For example, if teaching and learning are substitutes, $\alpha > 1$, and $b \to 1$ or $\beta_2 \to 1$ (take $\beta_2 \to 1$ and $b \in (0,1)$), then the simultaneous mode dominates

$$p\left(T^s_{i1}(c), L^s_{j1}(c)\right) < b\alpha p\left(T^n_{i1}(c,1,1), L^n_{j1}(c,1,1)\right) + (1-b)\alpha p\left(T^n_{i2}(c,1,1), L^n_{j2}(c,1,1)\right).$$

Examples relating to this case exhibit a declining marginal productivity of learning effort as teaching effort is increased. This might be true if learning entails the receiver working through applications on his or her own in order to assimilate the material. Further, the teacher's local adaptation to the needs of the student would play a major role in such applications oriented or "hands on" learning. Finally, the student would have a fairly low opportunity cost of learning effort in this case as would be the case with a trainee.

The sequential approach dominates the simultaneous approach in cases where teaching and learning are complements, where the local adaptation parameter is small and where the likelihood of high effort cost or the magnitude of high effort cost are great. For example, if teaching and learning are complements, $\beta_2 > 1$ and $\alpha = 1$, then the sequential mode dominates

$$p\left(T^s_{i1}(c), L^s_{j1}(c)\right) > bp\left(T^n_{i1}(c,1,1), L^n_{j1}(c,1,1)\right) + (1-b)p\left(T^n_{i2}(c,1,\beta_2), L^n_{j2}(c,1,\beta_2)\right)$$

Examples of this case would exhibit high opportunity cost of student effort and include ideas in which teacher effort enhances student marginal productivity of learning effort. Local adaptation of the teacher would not be important. Instructions on how to use software or equipment might fit this case in that a carefully crafted web presentation which anticipates all questions that the learner might have might be the best method of communication. In this case, the learner could be a busy executive with an associated high effort cost.

Our quadratic example allows additional comparative static analysis of changes in teaching and learning productivities. Let the basic technology be augmented as

$$\alpha p = \alpha(tT + \ell L - T^2 - L^2 + sTL),$$

where $\alpha = 1$ if the mode is sequential and $s \in \{-1, 1\}$. The parameters $t, \ell \geqslant 1$ are productivity parameters for teaching and learning. Each has the effect of increasing the teaching or learning marginal and average productivities at each T or L, respectively, all other things equal.

A change in teaching productivity yields

$$\frac{\partial \Omega^s(c)}{\partial t} = 2\pi(2t + s\ell)\frac{y-2c}{3}(1-\pi) \quad \text{and} \quad \frac{\partial \Omega^n(c)}{\partial t} = \alpha\frac{\partial \Omega^s(c)}{\partial t}.$$

Changes in learning productivity generate

$$\frac{\partial \Omega^s(c)}{\partial \ell} = 2\pi(2\ell + st)\frac{y-2c}{3}(1-\pi) \quad \text{and} \quad \frac{\partial \Omega^n(c)}{\partial \ell} \quad \text{and} \quad = \alpha\frac{\partial \Omega^s(c)}{\partial \ell}.$$

Conclusions for productivity changes. Taking $t = \ell = 1$ and $s \in \{-1, 1\}$, we have that increases in teaching or learning productivity increase the principal's expected profit at any c. However, the simultaneous mode results in greater increases in profit than the sequential mode with an increase in productivity, at a common c. This is driven by the presence of the local adaptation parameter in the simultaneous approach. Local adaptation enhances both learning and teaching productivity, so that it would lead naturally to a greater comparative static effect. However, the optimal c is always less in the dominant method with the parameters α, t, and ℓ close to 1. Thus, for α, t, and ℓ close to 1, we can say that profit is definitely more sensitive to productivity changes under the simultaneous approach if teaching and learning are substitutes. This is an interesting and clear prediction of our model brought out by our inclusion of an endogenously optimal contract. However, if teaching and learning are complements, then there is a trade off because the presence of small α makes the simultaneous approach more sensitive to changes in productivity, while the lower incentive payment under the sequential approach makes the sequential approach more sensitive to changes in productivity.

5. Two extensions: selling ideas to an external market and endogenizing local adaptation

5.1. Selling ideas to an external market

Suppose now that the idea developed within the firm can be marketed to an external set of students. The firm takes on the role of the teacher and potential external agents are the students. We assume that due to the prohibitive costs of establishing a physical classroom (or the technological superiority of the sequential mode), the firm adopts the sequential mode for knowledge sharing. Think of the number of potential students as the number of seats in a virtual classroom. Assume that students can have two different abilities and two different effort cost levels. This implies that for each of the N seats in the classroom, there are four different types of students that could occupy that seat as identified by learning efforts and abilities, namely

$$\{L_{11}, L_{12}, L_{21}, L_{22}\},$$

where L_{ij} denotes effort of a student with ability $i = 1, 2$ and effort cost $j = 1, 2$. Type 1 ability is high and type 2 is low, with associated probabilities a_1 and a_2 respectively. Type one effort cost is equal to $\beta_1 = 1$, while type 2 is $\beta_2 > \beta_1$, with associated probabilities b_1 and b_2, respectively. The number of possible classroom configurations of student types (hereafter called the number of classroom types) is the number of permutations of four types with repetition, N^4. For example, with two seats there are $2^4 = 16$ different classrooms each given by (L_{ij}, L_{lk}), $i, j, k, l = 1, 2$, where the classroom (L_{ij}, L_{lk}) materializes with the ex ante probability $a_i b_j a_l b_k$. In the example of N seats, one of the N^4 classrooms is given by $(L_{ij}, L_{lk}, \ldots L_{nr}) \in \mathbb{R}^n$ and such classroom has ex ante probability $a_i b_j a_l b_k \cdots a_n b_r$ (2 N terms) of materializing.

We assume that a student of ability i and effort cost j is willing to pay B_{ij} contingent on a knowledge transfer by the firm. The firm is a perfectly discriminating monopolist with respect to its knowledge or idea and charges B_{ij} contingent on information assimilation to a student. We assume that there is an acquisition cost of $G(N)$ associated with acquiring and N students and determining willingness to pay B_{ij} for each type, where the following assumption holds.

A.2 $G(0) = 0$, $\lim_{N\to 0} G(N) \geqslant 0$, G', $G'' > 0$, for $N \in (0, \infty)$.

Note that A.2 allows for a variable set up cost and that it assumes that acquisition cost is strictly convex on $(0, \infty)$. This assumption dictates that marginal acquisition cost is rising, but that average acquisition cost may be falling. The function p is assumed to take the general form $p_i(T, L_{ij})$ for a student of ability i.

For the purpose of developing intuition, take the case of a two seat classroom. Because high effort cost is avoided under the sequential mode, the teacher will exert a public good teaching effort T^s and face the following set of four possible classrooms

$$\{(L_{11}, L_{11}), (L_{21}, L_{21}), (L_{11}, L_{21}), (L_{21}, L_{11})\}.$$

Each student will obey his reaction function $L_{i1}(T^s)$ which is defined by the solution to

$$B_{i1} \frac{\partial p_i(T^s, L_{i1})}{\partial L_{i1}} - 1 = 0$$

With N fixed at 2, the teacher solves

$$\max_{\{T^s\}} \sum_{i=1}^{2} \sum_{j=1}^{2} a_i a_j [B_{i1} p_i(T^s, L_{i1}) + B_{j1} p_j(T^s, L_{j1}) - T^s] - G(2) \tag{22}$$

which can be rewritten as

$$\max_{\{T^s\}} 2[a_1 B_{11} p_i(T^s, L_{11}) + a_2 B_{21} p_j(T^s, L_{21})] - T^s - G(2). \tag{23}$$

Problem (23) simply scales the one person classroom problem above by doubling the benefit due to the addition of the second student, but the expression for teaching effort enters the expression the same because it is a public input.

For the case of N students to be chosen by the firm, (22) becomes

$$\max_{\{T^s, N\}} \sum_{i=1}^{2} \sum_{j=1}^{2} \cdots \sum_{k=1}^{2} a_i a_j \cdots a_k [B_{i1} p_i(T^s, L_{i1}) + B_{j1} p_j(T^s, L_{j1}) + \cdots + B_{k1} p_k(T^s, L_{k1})] - T^s - G(N)$$

which simplifies to

$$\max_{\{T^s, N\}} N[a_1 B_{11} p_1(T^s, L_{11}) + a_2 B_{21} p_2(T^s, L_{21})] - T^s - G(N). \tag{24}$$

The firm's optimal T^s is analogous to the one student case with a scale factor of N for the marginal benefit:

$$N\left\{ a_1 B_{11} \left[\frac{\partial p_1(T^s, L_{11})}{\partial T^s} + \frac{\partial p_1(T^s, L_{11})}{\partial L_{11}} \frac{\partial L_{11}^1}{\partial T} \right] + a_2 B_{21} \left[\frac{\partial p_2(T^s, L_{21})}{\partial T^s} + \frac{\partial p_2(T^s, L_{21})}{\partial L_{21}} \frac{\partial L_{21}^2}{\partial T^s} \right] \right\} - 1 = 0. \tag{25}$$

The optimal class size from the firm's view is described by

$$[a_1 B_{11} p_1(T^s, L_{11}) + a_2 B_{21} p_2(T^s, L_{21})] - G'(N) = 0. \tag{26}$$

Conclusions for external market. Ex ante expected profit per student is given by

$$\Pi/N \equiv a_1 B_{11} \alpha_1 p_1(T^s, L_{11}) + a_2 B_{21} p_2(T^s, L_{21}) - T^s/N - G(N)/N. \tag{27}$$

As N grows, the gross expected revenue per student is a constant and is the same as it is with a single student, and fixed teaching effort cost is spread over a larger number of students. Expected profit per student rises in N, if average acquisition cost falls. In the limit, we have that

$$\lim_{N\to\infty} \Pi/N = a_1 B_{11} p_1(T^s, L_{11}) + a_2 B_{21} p_2(T^s, L_{21}) - \lim_{N\to\infty} G(N)/N.$$

Thus, this "distance" mode of teaching has the beneficial scale effects popularly associated with it, if, first, the student acquisition cost function possesses variable set up costs so as to generate declining average acquisition cost and if, second, the limit of average acquisition cost as class size becomes arbitrarily large is less than gross expected revenue per student.

5.2. Endogenizing local adaptation

Let us retain the assumption that the sequential approach allows low effort cost self selection, because communication is asynchronous. The notion of local adaptation in the simultaneous method can be endogenized by assuming that students have different abilities for learning. The simultaneous method of knowledge sharing allows the teacher to locally adapt teaching effort to the ability level and the effort cost of the student at hand, whereas the sequential method requires the teacher to deliver one teaching effort to a student who could be of many possible ability levels. We assume that the simultaneous mode teacher observes both effort cost and ability during communication, whereas, while the sequential mode teacher knows that effort cost will be low, perceived ability is based on the distribution of abilities. Two questions of interest arise. First, does the sequential approach, with its focus on the average student, gain favor in situations where the average ability of the potential learner is higher? Second, does the simultaneous approach, with its ability to adapt to the student at hand, gain favor in cases where the variance in abilities is great?

Consider representative teaching and learning agents and the quadratic technology, A.1. Low effort cost will, as above, be given by 1 and high effort cost is $\beta_2 > 1$, with probabilities b and $(1-b)$, respectively. There are two learning ability levels parameterized by the productivity parameter for L. High ability (type one) is $\ell + \Delta$ and low ability is $\ell - \Delta$, each with probability 0.5. This formulation allows for a separation of Δ as a parameter which represents the variance in abilities and the parameter ℓwhich represents average ability.

For a pair consisting of a teaching agent and a learning agent, the basic technology under the sequential approach is

$$\begin{aligned} Ep^s = {} & .5\left(T^s + (\ell+\Delta)L^s_{11} - T^{s2} - L^{s2}_{11} + sT^sL^{s2}_{11}\right) \\ & + .5\left(T^s + (\ell-\Delta)L^s_{21} - T^{s2} - L^{s2}_{21} + sT^sL^{s2}_{21}\right), \end{aligned} \tag{28}$$

where T^s is common teaching effort and L^s_{i1} is learning effort of an agent with ability of type i and low effort cost, type 1. The technology for the simultaneous approach is

$$\begin{aligned} Ep^n = {} & .5\Big[b\left(T^n_{11} + (\ell+\Delta)L^n_{11} - T^{n2}_{11} - L^{n2}_{11} + sT^n_{11}L^{n2}_{11}\right) \\ & +(1-b)\left(T^n_{12} + (\ell+\Delta)L^n_{12} - T^{n2}_{12} - L^{n2}_{12} + sT^n_{12}L^{n2}_{12}\right)\Big] \\ & + .5\Big[b\left(T^n_{21} + (\ell-\Delta)L^n_{21} - T^{n2}_{21} - L^{n2}_{21} + sT^n_{21}L^{n2}_{21}\right) \\ & +(1-b)\left(T^n_{22} + (\ell-\Delta)L^n_{22} - T^{n2}_{22} - L^{n2}_{22} + sT^n_{22}L^{n2}_{22}\right)\Big], \end{aligned} \tag{29}$$

where T^n_{ij} denotes teaching effort of an agent with ability level i and effort cost j and L^n_{ij} denotes learning effort of an agent of ability i and effort cost j. The learning efforts under both methods solve the Nash problems analogous to those discussed above, and the teaching effort under the sequential approach solves $dEp^s/dT^s = 0$.

The effects of a change in Δ or ℓ, variance or average ability, respectively, on expected profit can be determined by considering the effects on the expression Ep^i, $i = s, n$. That is, sign $(\partial\Omega^i/\partial k)$=sign$(\partial Ep^i/\partial k)$, for $k = \Delta, \ell$ and $i = s, n$.

In the simultaneous approach, a change in the variance in abilities generates, in equilibrium,

$$\frac{\partial Ep^n}{\partial \Delta} = \frac{2\Delta}{3} > 0, \tag{30}$$

so that greater dispersion in abilities always leads to greater expected profit for the principal, if the simultaneous approach is utilized for communication. This result does not depend on whether teaching and learning are complements or substitutes and it holds at any c, including the optimal one. This is exactly what our intuition suggested in that the simultaneous approach counters greater variance in abilities with adaptation. Expression (30) and all of the other comparative statics of this subsection are derived in the appendix.

Under the sequential approach, in equilibrium,

$$\frac{\partial Ep^s}{\partial \Delta} = \frac{-s(c^s-1)}{3c^s(c^s+1)} + 0.5\Delta, \tag{31}$$

where c^s is the principal's optimum. The first term in (31) represents the indirect effects of a change in Δ through teaching effort and the second is the direct effect of the greater level of learning effort of the high ability type over the low type, given low effort cost. The sign of the first term in (31) depends on the magnitude of c^s in relation to unity. While a closed form solution for c^s is not possible, if we assume that the firm's cash flow arising from the implementation of knowledge, y, is not too small and that the probability that at least one agent will receive information, π, is not too close to unity, then $c^s > 1$.[10] Adopt these assumptions so that $c^s > 1$. Generally, a rise in Δ produces an increase in L^s_{11} and a decrease in L^s_{21}, but the latter effect is greater in absolute value than the former so that expected (average) effort goes down. If teaching and learning are substitutes, $s = -1$, then the decrease in average effort produces greater average teaching effort and the first term of (31) is strictly positive, making the entire term positive. A larger variance leads to the counter intuitive result that the principal's expected profit is increased, although teaching effort is not directly adaptable to each ability type. If teaching and learning are complements, then less average learning effort implies less average teaching effort and the first or indirect term of (31) is negative. In this case the two terms of (31) are opposite in sign and it is possible that greater variance does decrease the principal's expected profit, as intuition suggests. Whether this is true depends on the parameters. With complements ($s = 1$),

$$\frac{\partial Ep^s}{\partial \Delta} < 0 \quad \text{if} \quad \frac{2(c^s-1)}{3c^s(c^s+1)} > \Delta, \tag{32}$$

and this sign is necessarily negative for small variance as measured by Δ. Thus, the intuitive result that greater variance reduces the attractiveness of the sequential approach is only definitely true when variance is small and the two efforts are complements. The expression $2(c^s-1)/(3c^s(c^s+1))$ is maximized at $c^s = 2.41$ with a maximal value

[10] A sufficient condition for $c^s > 1$ is that $y > \frac{2.7916-1.6666\pi}{(1-\pi)}$.

of 0.114, and it tends to zero as c^s becomes infinite. Thus, if the cash flow prize y is large, making c^s large, or if $\Delta > 0.114$, then the intuitive result (32) is simply not true. We have

Conclusions for variance changes. Generally, a greater variance will increase expected profit under the simultaneous approach and it will do the same under the sequential, if teaching and learning are substitutes (with y sufficiently large). For these cases a relative comparison of the magnitudes of these effects is not possible, without more specific information. In the special case where variance in abilities is small and teaching and learning are complements, we obtain our intuitive result that greater variance raises expected profit in the simultaneous mode and lowers it in the sequential mode.

Next, consider changes in the average ability ℓ. In what follows, we normalize $\ell = 1$ and take $\Delta \in (0,0.5)$. The latter restriction on Δ is required to guarantee that $L^s_{21} > 0$ for $c > 1$ and $s = -1$, in the sequential mode of communication. Variations in average ability produce the following changes in the simultaneous approach

$$\frac{\partial Ep^n}{\partial \ell} = [b + (1-b)\beta]\frac{s+2}{3c^n} + .5\left[(b(L^n_{11} + L^n_{21}) + (1-b)(L^n_{12} + L^n_{22})\right] > 0, \quad \text{for } s \in \{-1, 1\}. \tag{33}$$

The conclusion is that increases in average ability unambiguously raise the principal's expected profit regardless of the complement-substitute relationship between teaching and learning.

The sequential mode produces the following effect

$$\frac{\partial Ep^s}{\partial \ell} = \frac{s}{3} + 0.5(L^s_{11} + L^s_{21}). \tag{34}$$

The second term of (34) is the direct effect of an increase in average ability through learning efforts and this is always positive. The first term is indirect effect of a change in average ability through T^s and this effect is positive for complements and negative for substitutes. The explanation is that an increase in average ability increases both L^s_{11} and L^s_{21}. If teaching and learning are complements, then this increase increases common teaching effort and conversely if they are substitutes. In the complement case, increases in average ability always increase expected profit in equilibrium. In the substitute case, there are opposing signs and (34) can be further specialized to

$$\frac{\partial Ep^s}{\partial \ell} = \frac{-1}{6c^s(c^s+1)}(3.0c^s - 1.0c^s\Delta + c^{s2}\Delta - 2.0c^{s2} + 1.0)$$

which carries the sign of

$$c^{s2}(2-\Delta) - c^s(3+\Delta) - 1. \tag{35}$$

Given $\Delta \in (0,0.5)$ and $c^s > 1$, it is theoretically possible for (35) to be negative as long as the optimal c is sufficiently small. For example, $c^{s2}(2-\Delta) - c^s(3+\Delta) - 1 < 2c^{s2} - 3c^s - 1$ and the latter is negative and feasible if $c^s \in (0, 1.78)$. This range of optimal c is possible if y is adjusted downward such that $y/2 < (1.78)$.[11] It is then possible for an increase in average ability to result in less expected profit for the principal if the firm's reward for successful communication, y, is relatively small, if teaching and learning are substitutes, and if the sequential mode of communication is being used. This is an interesting and counter intuitive result.

Conclusions for average ability changes. Greater average ability raises expected profit under both approaches and a relative magnitude comparison of profit changes cannot be made without more specific information. In the special case where the firm's reward, y, for successful communication is small and teaching and learning are substitutes, a greater average ability actually decreases expected profit in the sequential approach and increases it under the simultaneous.

6. Conclusion

When knowledge transfer involves fairly complex ideas, it can be characterized as a joint production process requiring that both the sender and the receiver exert effort. Production of both the communication and assimilation of an idea then is a team process where each member incurs a private cost but may not receive the full public benefit. If communication is necessary to produce additional cash flow within an organization, a principal can choose the mode of knowledge transfer, sequential or simultaneous, and incentivize such costly transfer by properly rewarding cash flow. The optimal contract rewards cash flow only if the sending and receiving agents successfully complete the transfer and generate the additional cash flow that the communicated idea can make possible. At the optimal contract, there is under supply of both teaching and learning efforts in both the sequential and the simultaneous modes, regardless of whether teaching and learning are substitutes or complements. The optimal reward for successful transfer varies inversely with the probability that an agent will be endowed with a new idea for communication and the principal's equilibrium expected profit varies directly with this probability.

A key result is that in fairly general circumstances, the sequential mode generates a greater expected profit for the principal and it requires a lower optimal reward for communication if teaching and learning efforts are complements. On the other hand, if teaching and learning are substitutes, the simultaneous mode produces greater equilibrium profit and requires less payment for successful knowledge transfer. This result points out a key difference between the two modes in a situation where the technologies are identical except for their sequentiality or simultaneity. The endogeneity of the optimal reward for knowledge transfer adds the result that the dominant method of communication requires a lower incentive payment than the alternative method.

There are differences between the two approaches to knowledge transfer which are not described in the basic model. The sequential mode has the advantage of allowing receivers to access information when their effort costs are low and the simultaneous mode allows the teacher to locally adapt to the student's ability and effort cost.

[11] This is true because at $c = y/2$, $\Omega^{s\prime}(c) = -2\pi^2 - 4(1-\pi)\pi Ep^s < 0$.

Modifying the basic model to account for these features, we show that the simultaneous approach dominates the sequential approach in situations where teaching and learning are substitutes, where the local adaptation parameter is large, where the likelihood that the student will have low effort cost is high, or where the magnitude of the high effort cost is relatively low. The sequential approach dominates the simultaneous approach in cases where teaching and learning are complements, where the local adaptation parameter is small and where the likelihood of high effort cost or the magnitude of high effort cost are great.

Our comparative statics show that increases in the productivity of teaching or learning will raise expected profit at a given payment for successful communication under both approaches. We find that when teaching and learning are substitutes, the simultaneous approach generates greater profit gains than the sequential if teaching or learning productivities are increased at optimal contracts. If teaching and learning are complements, there is a trade off in that the sequential mode has a lower incentive payment but the simultaneous mode has local adaptation. Here the comparative magnitudes of changes brought by productivity increases are ambiguous.

We examined the external marketing of the knowledge developed in the firm through the use of the sequential mode. This extension illustrates that the sequential mode of communication has beneficial scale effects if acquisition costs per student decline as class increases, due to variable set up costs, and if the limit of average acquisition cost as class size becomes arbitrarily large is less than gross expected revenue per student. Thus, the often touted scale benefits of "distance learning" are present only under some specific qualifying conditions.

We endogenize the local adaptation aspect of the simultaneous approach by assuming that potential students can be of either high or low ability and that simultaneous communication can adapt to either effort cost or ability level. In this version of the model, we study the effects of increases in the average ability and increases in the variance of abilities on the principal's expected profit. Under the sequential mode, a rise in the variance of abilities unexpectedly raises expected profit, if teaching and learning are substitutes (and y is sufficiently large), but it leads to an ambiguous effect if they are complements in that expected profit can go up or down. It decreases expected profit with complements in the special case where the variance is small. Under the simultaneous mode, a greater dispersion in abilities always leads to greater expected profit for the principal. Only in the special case where the variance in abilities is small and teaching and learning are complements do we get our conjectured result that greater variance definitely favors the simultaneous approach over the sequential. Increases in average ability increase expected profit if the simultaneous mode is used regardless of whether teaching and learning are complements or substitutes. In the sequential mode an increase in average ability raises expected profit if teaching and learning are complements. However if they are substitutes in the sequential approach, a rise in average ability can actually lead to less expected profit, if the prize for successful application of a communicated idea is small. It is only in this case that we can say that an increase in average ability definitely favors one approach versus the other and the favored approach is the simultaneous one.

Acknowledgements

I gratefully acknowledge the generous financial support provided by the Marshall School of Business. I benefited from discussions with Ricardo Alonso, Odilon Camera, Richard Eastin, Krishna Kumar, Oguzhan Ozbas, Michael Raith, Heikki Rantakari, Jan Zabojnik and Fernando Zapatero.

Appendix A

The principal's problem for the simultaneous move case of Section 2

The principal's problem is (Lagrangian multipliers are $(\mu_i, \lambda_i, \gamma_i)$)

$$\begin{aligned}\max_{\{S,c,T_i,L_j\}} L &= 2(1-\pi)\pi y + 2\pi^2(y-c) + (1-\pi)\pi p(T_1,L_2)(y-2c)\\ &\quad + (1-\pi)\pi p(T_2,L_1)(y-2c)\\ &\quad - 2S + \mu_1[c\partial p(T_1,L_2)/\partial T_1 - 1] + \mu_2[c\partial p(T_2,L_1)/\partial L_1 - 1]\\ &\quad + \mu_3[c\partial p(T_2,L_1)/\partial T_2 - 1] + \mu_4[c\partial p(T_1,L_2)/\partial L_2 - 1]\\ &\quad + \sum_{i=1}^{2} \lambda_i\{S + c[\pi^2 + \pi(1-\pi)[p(T_i,L_j) + p(T_j,L_i)]] - T_i - L_i\}\\ &\quad + \gamma_c c + \gamma_S S.\end{aligned}$$

Consider the first order conditions. Assume that the participation constraints are non-binding, that the incentive compatibility constraints are binding at an interior solution. The first order condition for S is given by

$$-2 + \lambda_1 + \lambda_2 + \gamma_S = 0.$$

If the participation constraints are nonbinding, then $\lambda_i = 0$, $i = 1, 2$, and $\gamma_S > 0$. Whence the limited liability constraint for S is binding and $S = 0$.

The principal's problem for the sequential move case of Section 2

The principal's problem is (Lagrangian multipliers are $(\mu_i, \lambda_i, \gamma_i)$)

$$\begin{aligned}\max_{\{c,S,T_i\}} L &= 2(1-\pi)\pi y + 2\pi^2(y-c)\\ &\quad + (1-\pi)\pi p(T_1, L_2(T_1,c))(y-2c)\\ &\quad + (1-\pi)\pi p(T_2, L_1(T_2,c))(y-2c) - 2S\\ &\quad + \mu_1\{c[\partial p(T_1, L_2(T_1,c))/\partial T_1\\ &\quad + (\partial L_2(T_1,c)/\partial T_1)\partial p(T_1, L_2(T_1,c))/\partial L_2] - 1\}\\ &\quad + \mu_2\{c[\partial p(T_2, L_1(T_2,c))/\partial T_2\\ &\quad + (\partial L_1(T_2,c)/\partial T_2)\partial p(T_2, L_1(T_2,c))/\partial L_1] - 1\}\\ &\quad + \sum_{i\neq j=1}^{2} \lambda_i[S + c\{\pi^2 + \pi(1-\pi)[p(T_i, L_j(T_i,c))\\ &\quad + p(T_j, L_i(T_j,c))]\} - T_i - L_i(T_j,c)] + \gamma_c c + \gamma_S S.\end{aligned}$$

The (FOC) for S is given by

$$-2 + \lambda_1 + \lambda_2 + \gamma_S = 0.$$

If we assume that the participation constraints are nonbinding and that there is an interior solution (we will give sufficiency conditions later), then $\lambda_i = 0$, $\gamma_s > 0$ and $S = 0$.

Proof of Lemma 1. For the existence of a $c \in (1, y/2) > 1$ such that $\Omega^{n\prime}(c) = 0$, it suffices that

$$\lim_{c \to 1} \Omega^{n\prime}(1) > 0 \quad \text{and} \quad \lim_{c \to y/2} \Omega^{n\prime}(c) < 0.$$

First consider the case of complements. We have $\Omega^{n\prime}(c) = 2\frac{\pi}{c^3}(2y - 2c - 2\pi y + \pi c^3 + 2\pi c - 2c^3)$ and $\Omega^{n\prime}(1) = 2\pi(3\pi + 2y(1 - \pi) - 4) > 0$ if $y > \frac{4-3\pi}{2(1-\pi)}$. Given $\pi \in (0,1)$, the latter condition is implied by $y > \frac{4-\pi}{2(1-\pi)}$. We have

$$\Omega^{n\prime}\left(\frac{y}{2}\right) = 2\frac{\pi}{c^3}\left(y(1-\pi) + \frac{1}{8}\pi y^3 - \frac{1}{4}y^3\right) < 0$$
$$\text{if } y > \left(\frac{8(1-\pi)}{(2-\pi)}\right)^{1/2}.$$

However, for $\pi \in (0,1)$, $\frac{4-\pi}{2(1-\pi)} > \left(\frac{8(1-\pi)}{(2-\pi)}\right)^{1/2}$, so that the result holds. For the case of substitutes,

$$\Omega^{n\prime}(c) = -\frac{2}{3}\frac{\pi}{c^3}(2c - 2y + \pi c^3 - 2\pi c + 2\pi y + 2c^3) \quad \text{and}$$
$$\Omega^{n\prime}(1) = \frac{2}{3}\pi(\pi + 2y - 2\pi y - 4) > 0 \quad \text{if } y > \frac{4-\pi}{2(1-\pi)},$$

which is true. Next,

$$\Omega^{n\prime}\left(\frac{y}{2}\right) = -\frac{16}{3}\frac{\pi}{y^3}\left(\frac{1}{8}\pi y^3 - y + \pi y + \frac{1}{4}y^3\right)$$
$$< 0 \quad \text{if } y > \left(\frac{8(1-\pi)}{\pi+2}\right)^{1/2} < \frac{4-\pi}{2(1-\pi)},$$

and the result holds. Given that we have shown that there exists a $c \in (1, y/2)$ such that $\Omega^{n\prime}(c) = 0$, for a unique solution it suffices to show that $\Omega^{n\prime}(c) < 0$. For complements, $\Omega^{n\prime\prime}(c) = 4\frac{\pi}{c^4}(2c - 3y)(1 - \pi) < 0$, by $\pi \in (0,1)$ and $c < y/2$. For substitutes, $\Omega^{n\prime\prime}(c) = \frac{4}{3}\frac{\pi}{c^4}(2c - 3y)(1 - \pi) < 0$, by the same assumptions. To show that optimal expected profit is positive $\Omega^s(y/2)|_{s\in\{-1,1\}} = \pi y(2 - \pi) > 0$. □

Proof of Proposition 1. The text presents all of the results except for the comparative statics of π and c. It is more efficient to proceed with general functional notation so as to cover the cases of complements and substitutes in one proof. The sign $\partial c/\partial \pi = \text{sign}[-2\pi + (1 - 2\pi)p^{n\prime}(c)(y - 2c) - 2(1 - 2\pi)p^n(c)]$. Substituting from the first order condition, the latter expression can be written as

$$\text{sign}[-2\pi + (1-2\pi)(\pi/(1-\pi) + 2p^n(c)) - 2(1-2\pi)p^n(c)]$$
$$= \text{sign}\left[-2 + \frac{1-2\pi}{1-\pi}\right] < 0.$$

Thus, $\partial c/\partial \pi < 0$. Next, consider $\partial \Omega^n(c)/\partial \pi = 2(1 - 2\pi)y + 4\pi(y - c) + 2(1 - 2\pi)p^n(c)(y - 2c)$. Clearly if $\pi \leqslant 0.5$, then $\partial \Omega^n(c)/\partial \pi > 0$. If $\pi > 0.5$, then $\partial \Omega^n(c)/\partial \pi \gtreqless 0$ as $0 \gtreqless -(y - 2c\pi) + p^n(c)(y - 2c)(2\pi - 1)$. The

$$\lim_{\pi \to 1} -(y - 2c\pi) + p^n(c)(y - 2c)(2\pi - 1)$$
$$= -(y - 2c) + p^n(c)(y - 2c) < 0, \text{ and}$$

$$\lim_{\pi \to 0.5} -(y - 2c\pi) + p^n(c)(y - 2c)(2\pi - 1) = -(y - c) < 0.$$

Further,

$$\frac{\partial}{\partial \pi}[-(y - 2c\pi) + p^n(c)(y - 2c)(2\pi - 1)]$$
$$= 2c + 2p^n(c)(y - 2c) > 0.$$

It follows that $\partial \Omega^n(c)/\partial \pi > 0$. □

Proof of Lemma 2. For the existence of a $c \in (1, y/2) > 1$ such that $\Omega^{s\prime}(c) = 0$, it suffices that $\lim_{c\to 1}\Omega^{s\prime}(1) > 0$ and $\lim_{c \to y/2}\Omega^{s\prime}(c) < 0$. First consider the case of complements. We have

$$\Omega^{s\prime}(c) = -\frac{1}{3}\frac{\pi}{c^3}(7c - 7y - 6\pi c^3 - 7\pi c + 7\pi y + 12c^3)$$

and

$$\Omega^{s\prime}(1) = \frac{1}{3}\pi(13\pi + 7y - 7\pi y - 19) > 0$$

if $y > \frac{19-13\pi}{7(1-\pi)}$. We have

$$\Omega^{s\prime}\left(\frac{y}{2}\right) = -\frac{8}{3}\frac{\pi}{y^3}\left(\frac{7}{2}\pi y - \frac{3}{4}\pi y^3 - \frac{7}{2}y + \frac{3}{2}y^3\right) < 0$$

if $y > \left(\frac{\frac{7}{2}(1-\pi)}{\left(\frac{3}{2}-\frac{3}{4}\pi\right)}\right)^{1/2}$. However, for $\pi \in (0,1), \frac{19-13\pi}{7(1-\pi)} > \left(\frac{\frac{7}{2}(1-\pi)}{\left(\frac{3}{2}-\frac{3}{4}\pi\right)}\right)^{1/2}$, so that the result holds. For the case of substitutes,

$$\Omega^{s\prime}(c) = -\frac{1}{3}\frac{\pi}{c^3}(7c - 7y + 2\pi c^3 - 7\pi c + 7\pi y + 4c^3)$$

and

$$\Omega^{s\prime}(1) = \frac{1}{3}\pi(5\pi + 7y - 7\pi y - 11) > 0$$

if $y > \frac{11-5\pi}{7(1-\pi)}$, which is true by $\frac{19-13\pi}{7(1-\pi)} > \frac{11-5\pi}{7(1-\pi)}$, for $\pi \in (0,1)$. Next,

$$\Omega^{s\prime}\left(\frac{y}{2}\right) = -\frac{8}{3}\frac{\pi}{y^3}\left(\frac{1}{4}\pi y^3 - \frac{7}{2}y + \frac{7}{2}\pi y + \frac{1}{2}y^3\right) < 0$$

if $y > \left(\frac{7(1-\pi)}{(1+.5\pi)}\right)^{1/2} < \frac{19-13\pi}{7(1-\pi)}$, and the result holds. Given that we have shown that there exists a $c \in (1, y/2)$ such that $\Omega^{n\prime}(c) = 0$, for a unique solution, it suffices to show that $\Omega^{n\prime}(c) < 0$. For complements, $\Omega^{s\prime\prime}(c) = \frac{7}{3}\frac{\pi}{c^4}(2c - 3y)(1 - \pi) < 0$ which is true if $c < \frac{3}{2}y$, by $\pi \in (0,1)$. Given $c < y/2$, the result holds. For substitutes, $\Omega^{s\prime\prime}(c) = \frac{7}{3}\frac{\pi}{c^4}(2c - 3y)1 - \pi) < 0$, by the same assumptions. To show that optimal expected profit is positive it suffices to show that $\Omega^s(y/2)|_{s\in\{-1,1\}} = \pi y(2 - \pi) > 0$. □

Proof of Proposition 2. The text provides all of the results except for the comparative statics of π and c. It is more efficient to proceed with general functional notation so as to cover the cases of complements and substitutes in one proof. The sign$\partial c/\partial \pi = \text{sign}[-2\pi + (1 - 2\pi)p^{s\prime}(c)(y - 2c) - 2(1 - 2\pi)p^s(c)]$. Substituting from the first order condition, the latter expression can be written as

$$\mathrm{sign}[-2\pi + (1-2\pi)(\pi/(1-\pi) + 2p^s(c)) - 2(1-2\pi)p^s(c)] = \mathrm{sign}\left[-2 + \frac{1-2\pi}{1-\pi}\right] < 0.$$

Thus, $\partial c/\partial\pi < 0$. Next, consider $\partial\Omega^s(c)/\partial\pi = 2(1-2\pi)y + 4\pi(y-c) + 2(1-2\pi)p^s(c)(y^* - 2c)$. Clearly if $\pi \leqslant 0.5$, then $\partial\Omega^s(c)/\partial\pi > 0$. If $\pi > 0.5$, then $\partial\Omega^s(c)/\partial\pi \gtreqless 0$ as $0 \gtreqless -(y-2c\pi) + p^s(c)(y-2c)(2\pi-1)$. The

$$\begin{aligned}\lim_{\pi\to 1} & -(y-2c\pi) + p^s(c)(y-2c)(2\pi-1) \\ &= -(y-2c) + p^s(c)(y-2c) \\ &< 0, \text{ and } \lim_{\pi\to .5} -(y-2c\pi) + p^s(c)(y-2c)(2\pi-1) \\ &= -(y-c) < 0.\end{aligned}$$

Further,

$$\begin{aligned}&\frac{\partial}{\partial\pi}[-(y-2c\pi) + p^s(c)(y-2c)(2\pi-1)] \\ &= 2c + 2p^s(c)(y-2c) > 0.\end{aligned}$$

It follows that $\partial\Omega^s(c)/\partial\pi > 0$. □

Proof of Proposition 3

For the case of complements

$$p^s(c) - p^n(c) = \frac{5}{12c^2} > 0.$$

At either solution the principal's objective function can be written as

$$\Omega^i = 2\pi^2(y-c) + 2\pi(1-\pi)y + 2\pi(1-\pi)p^i(c)(y-2c),$$
$i = s, n$.

Given that $(y-2c) > 0$, we have

$$\begin{aligned}\Omega^s(c) &= 2\pi^2(y-c) + 2\pi(1-\pi)y + 2\pi(1-\pi)p^s(c)(y-2c) \\ &> 2\pi^2(y-c) + 2\pi(1-\pi)y + 2\pi(1-\pi)p^n(c)(y-2c) \\ &= \Omega^n(c), \text{ for all } c,\end{aligned}$$

because $p^s(c) > p^n(c)$, for all $c < y/2$. It follows that the sequential move dominates.

To show that $c^n > c^s$, it suffices to show that

$$\Omega^{s\prime}(c) - \Omega^{n\prime}(c) = -\frac{5}{3}\frac{\pi}{c^3}(y-c)(1-\pi) < 0.$$

This is true by $(y-c)(1-\pi) > 0$.

For the case of substitutes

$$p^n(c) - p^s(c) = \frac{1}{4c^2} > 0$$

From the case of complements, we have that $\Omega^s(c) < \Omega^n(c)$ if $p^s(c) < p^n(c)$, for all $c < y/2$.

To show that $c^s > c^n$, it suffices to show that

$$\Omega^{n\prime}(c) - \Omega^{s\prime}(c) = -\frac{\pi}{c^3}(y-c)1-\pi) < 0,$$

which is true by $(y-c)(1-\pi) > 0$. □

Proof of Proposition 4. Consider the complements case first. In the simultaneous move regime, a teaching agent's incentive compatibility (*IC*) condition says

$$c[\partial p(T_i, L_j)/\partial T_i] - 1 = 0.$$

A teaching agent's (*IC*) condition for the sequential move case is

$$c[\partial p(T_i, L_j(T_i, c))/\partial T_i + (\partial L_j(T_i, c)/\partial T_i)(\partial p(T_i, L_j(T_i, c))/\partial L_j)] - 1 = 0.$$

Evaluate the latter (*IC*) at T_i^n, where we have that $c\left[\partial p\left(T_i^n, L_j^n\right)/\partial T_i\right] - 1 = 0,$

$$1 + \left(\partial L_j\left(T_i^n, c\right)/\partial T_i\right)(1) - 1 = \partial L_j\left(T_i^n, c\right)/\partial T_i > 0.$$

Let $\omega_i^T\left(T_i^n, c\right) = cp(T_i, L_j(T_i, c)) - T_i$. We have that $\partial\omega_i^T\left(T_i^n, c\right)/\partial T_i > 0$, while $\partial\omega_i^T\left(T_i^s, c\right)/\partial T_i = 0$, with $\partial^2\omega_i^T\left(T_i^n, c\right)/\partial T_i^2 = c\frac{\partial^2 p(T_i, L_j(T_i, c))}{\partial T_i^2} < 0$. Given that $\partial\omega_i^T\left(T_i^n, c\right)/\partial T_i$ is strictly decreasing in T_i, it follows that $T_i^s > T_i^n$. By $L_j(T_i, c)$ strictly increasing in T_i, it follows that $L_j\left(T_i^s, c\right) > L_j\left(T_i^n, c\right)$. We have shown that teaching and learning efforts are greater in the sequential equilibrium at any given c. Whence, $p^s(c) - p^n(c) > 0$ and the sequential mode dominates.

In the case of substitutes, again write the teaching agent's (*IC*) at the sequential move equilibrium evaluated at the simultaneous move effort levels

$$\begin{aligned}\partial\omega_i^T\left(T_i^n, c\right)/\partial T_i &= 1 + \left(\partial L_j\left(T_i^n, c\right)/\partial T_i\right)(1) - 1 \\ &= \partial L_j\left(T_i^n, c\right)/\partial T_i < 0.\end{aligned}$$

Thus, $\partial\omega_i^T\left(T_i^n, c\right)/\partial T_i < 0$, while $\partial\omega_i^T\left(T_i^s, c\right)/\partial T_i = 0$, with $\partial^2\omega_i^T(T_i, c)/\partial T_i^2 = c\frac{\partial^2 p(T_i, L_j(T_i, c))}{\partial T_i^2} < 0$. It follows that $T_i^s < T_i^n$. By $L_j(T_i, c)$ strictly decreasing in T_i, it follows that $L_j\left(T_i^s, c\right) > L_j\left(T_i^n, c\right)$. We then have that $T_i^s < T_i^n$, while $L_j^s > L_j^n$. Note that the assumption $|\frac{p_{ii}}{p_i}| > |\frac{p_{ij}}{p_j}|, i \neq j$, implies that $p(T_i, L_j(T_i, c))$ is strictly increasing in T_i. That is, sign $(\partial p(T_i, L_j(T_i, c))/\partial T_i) = \mathrm{sign}\left(\frac{-\partial^2 p(T_i, L_j)/\partial L_j^2}{\partial p(T_i, L_j)/\partial L_j} + \frac{\partial^2 p(T_i, L_j)/\partial L_j\partial T_i}{\partial p(T_i, L_j)/\partial T_i}\right) > 0$. It follows that $p^s(c) = p\left(T_i^s, L_j\left(T_i^s, c\right)\right) < p\left(T_i^n, L_j^n\left(T_i^n, c\right)\right) = p^n(c)$ and the simultaneous mode dominates. □

Proof of Lemma 3

Let J denote the Jacobian of the system (16) and (17). We have

$$\begin{aligned}|J| &= c^2\alpha^2 p_{TT}\left(T_{ik}^n, L_{jk}^n\right)p_{LL}\left(T_{ik}^n, L_{jk}^n\right) - c^2\alpha^2\left[p_{TL}\left(T_{ik}^n, L_{jk}^n\right)\right]^2 \\ &= c^2\alpha^2\left|H\left(T_{ik}^n, L_{jk}^n\right)\right| > 0,\end{aligned}$$

where H is the Hessian of p. Consider α, so that

$$\begin{aligned}\partial T_{ik}^n/\partial\alpha = [1/|J|]\Big[&-c\alpha p_{LL}\left(T_{ik}^n, L_{jk}^n\right)cp_T\left(T_{ik}^n, L_{jk}^n\right) \\ &+ c\alpha p_{TL}\left(T_{ik}^n, L_{jk}^n\right)cp_L\left(T_{ik}^n, L_{jk}^n\right)\Big] > 0.\end{aligned}$$

The proof of $\partial L_{jk}^n/\partial\alpha > 0$ is analogous, and $\partial p\left(T_{ik}^n, L_{jk}^n\right)/\partial\alpha > 0$ directly follows from these results and $p_i > 0$.

We have $\partial T^n_{i2}/\partial\beta_2 = [1/|J|]\left[-c\alpha p_{TL}\left(T^n_{i2}, L^n_{j2}\right)\right] < 0$ if complements and > if substitutes. Further, $\partial L^n_{j2}/\partial\beta_2 = [1/|J|][c\alpha p_{TT}\left(T^n_{i2}, L^n_{j2}\right)] < 0$. The derivative $\partial p\left(T^n_{i2}, L^n_{j2}\right)/\partial\beta_2 = p_T\left(T^n_{i2}, L^n_{j2}\right)\partial T^n_{i2}/\partial\beta_2 + p_L\left(T^n_{i2}, L^n_{j2}\right)\partial L^n_{j2}/\partial\beta_2$ so that its sign is that of

$$[1/|J|][c\alpha]\left\{p_T\left(T^n_{i2}, L^n_{j2}\right)\left[-p_{TL}\left(T^n_{i2}, L^n_{j2}\right)\right] + p_L\left(T^n_{i2}, L^n_{j2}\right)\left[p_{TT}\left(T^n_{i2}, L^n_{j2}\right)\right]\right\}$$

which is negative. □

The principal's problem for the simultaneous move case of Section 4

The principal's problem is (Lagrangian multipliers are $(\mu_{ik}, \lambda_i, \gamma_i)$)

$$\begin{aligned}
\max_{\{S,c,T_{ik},L_{jk}\}} L = {} & 2(1-\pi)\pi y + 2\pi^2(y-c) + (1-\pi)\pi\left[b\alpha p\left(T^n_{11}, L^n_{21}\right)\right. \\
& \left.+ (1-b)\alpha p\left(T^n_{12}, L^n_{22}\right)\right](y-2c) \\
& + (1-\pi)\pi\left[b\alpha p\left(T^n_{21}, L^n_{11}\right) + (1-b)\alpha p\left(T^n_{22}, L^n_{12}\right)\right] \\
& \times (y-2c) - 2S + \sum_{k=1}^{2}\mu_{1k}\left[c\alpha\partial p\left(T^n_{1k}, L^n_{2k}\right)/\partial T_{1k} - 1\right] \\
& + \sum_{k=1}^{2}\mu_{2k}\left[c\alpha\partial p\left(T^n_{2k}, L^n_{1k}\right)/\partial L_{1k} - \beta_k\right] \\
& + \sum_{k=1}^{2}\mu_{3k}\left[c\alpha\partial p\left(T^n_{2k}, L^n_{1k}\right)/\partial T_{2k} - 1\right] \\
& + \sum_{k=1}^{2}\mu_{4k}\left[c\alpha\partial p\left(T^n_{1k}, L^n_{2k}\right)/\partial L_{2k} - \beta_k\right] \\
& + \lambda_1\{S + c[\pi^2 + \pi(1-\pi)\left[b\alpha p\left(T^n_{11}, L^n_{21}\right)\right. \\
& + (1-b)\alpha p\left(T^n_{12}, L^n_{22}\right) + b\alpha p\left(T^n_{21}, L^n_{11}\right) \\
& \left.+ (1-b)\alpha p\left(T^n_{22}, L^n_{12}\right)\right] - bT^n_{11} - (1-b)T^n_{12} \\
& + bL^n_{11} + (1-b)L^n_{12}\} + \lambda_2\{S + c[\pi^2 \\
& + \pi(1-\pi)\left[b\alpha p\left(T^n_{21}, L^n_{11}\right) + (1-b)\alpha p\left(T^n_{22}, L^n_{12}\right)\right. \\
& \left.+ b\alpha p\left(T^n_{11}, L^n_{21}\right) + (1-b)\alpha p\left(T^n_{12}, L^n_{22}\right)\right] \\
& - bT^n_{21} - (1-b)T^n_{22} + bL^n_{21} + (1-b)L^n_{22}\} + \gamma_c c + \gamma_S S.
\end{aligned}$$

The solution to this problem is analogous to the case where there is a single effort cost, if we assume that the participation constraints are nonbinding and that the incentive compatibility constraints are binding. The first order condition for S is again given by

$$-2 + \lambda_1 + \lambda_2 + \gamma_S = 0.$$

If the participation constraints are nonbinding, it then follows that $\lambda_i = 0$, $i = 1, 2$, and $\gamma_S > 0$. The optimal S is again zero, as the limited liability constraint for S is binding.

Proof of Lemma 4. For existence of a $c \in (0,y/2)$ such that $\Omega^{n\prime}(c) = 0$, it suffices to show that $\Omega^{n\prime}(0)|_{s\in\{-1,1\}} > 0$ and $\Omega^{n\prime}(y/2)|_{s\in\{-1,1\}} < 0$. For $s = 1$,

$$\begin{aligned}
\Omega^{n\prime}(0)|_{s=1} = \lim_{c\to 0}\Bigg[-\frac{4}{3}\frac{\pi}{c^3}\frac{y}{\alpha^2}(\pi-1)(2b + \beta_2(1-b) \\
+\beta_2^2(1-b) + 1)\Bigg] = \infty.
\end{aligned}$$

For $s = 1$,

$$\begin{aligned}
\Omega^{n\prime}(y/2)|_{s=1} = \frac{16}{3}\frac{\pi}{y^3\alpha^2}\Bigg(& y - \frac{3}{4}y^3\alpha^2 + y\beta_2 + y\beta_2^2 - \pi y + 2by \\
& + \frac{3}{8}\pi y^3\alpha^2 - \pi y\beta_2 - by\beta_2 - \pi y\beta_2^2 - by\beta_2^2 \\
& - 2\pi by + \pi by\beta_2 + \pi by\beta_2^2\Bigg),
\end{aligned}$$

so that $\Omega^{n\prime}(y/2)|_{s=1} < 0$ iff $y > \left(\frac{8}{3\alpha^2}\frac{1-\pi}{2-\pi}(2b + \beta_2 - b\beta_2 + \beta_2^2 - b\beta_2^2 + 1)\right)^{1/2}$. For $s = -1$,

$$\begin{aligned}
\Omega^{n\prime}(0)|_{s=-1} &= \lim_{c\to 0}\left[-\frac{4}{3}\frac{\pi}{c^3}\frac{y}{\alpha^2}(\pi-1)(\beta_2((1-b)(\beta_2-1))+1)\right] \\
&= \infty.
\end{aligned}$$

For $s = -1$,

$$\begin{aligned}
\Omega^{n\prime}(y/2)|_{s=-1} = -\frac{16}{3}\frac{\pi}{y^3\alpha^2}\Bigg(& \frac{1}{4}y^3\alpha^2 - y + y\beta_2 - y\beta_2^2 + \pi y \\
& + \frac{1}{8}\pi y^3\alpha^2 - \pi y\beta_2 - by\beta_2 + \pi y\beta_2^2 + by\beta_2^2 \\
& + \pi by\beta_2 - \pi by\beta_2^2\Bigg),
\end{aligned}$$

so that $\Omega^{n\prime}(y/2)|_{s=-1} < 0$ iff $y > \left(\frac{8}{\alpha^2(2+\pi)}((1-\pi)(1+\beta_2(\beta_2-1)(1+b)))\right)^{1/2}$. Thus, there exists a $c \in (0,y/2)$ such that $\Omega^{n\prime}(c) = 0$. To complete the proof, we need only show that $\Omega^{n\prime\prime}(c) < 0$. The sign of the latter is determined by the expression

$$2\pi(1-\pi)Ep^n(c,\alpha,\beta_2)(y-2c),$$

where $Ep^n(c,\alpha,\beta_2)$ is a convex combination of

$$\alpha p\left(T^n_{i1}(c,\alpha,1), L^n_{j1}(c,\alpha,1)\right) \text{ and } \alpha p\left(T^n_{i2}(c,\alpha,\beta_2), L^n_{j2}(c,\alpha,\beta_2)\right).$$

Our previous analysis shows that

$$\frac{\partial^2}{\partial c^2}\left(b2\pi(1-\pi)\alpha p\left(T^n_{i1}(c,\alpha,1), L^n_{j1}(c,\alpha,1)\right)(y-2c)\right) < 0.$$

We need only show that

$$\frac{\partial^2}{\partial c^2}\left(b2\pi(1-\pi)\alpha p\left(T^n_{i2}(c,\alpha,\beta_2), L^n_{j2}(c,\alpha,\beta_2)\right)(y-2c)\right) < 0. \tag{$*$}$$

For $s = 1$, condition $(*)$ can be expressed as

$$-\frac{4}{3}\frac{\pi}{c^4\alpha^2}(2c-3y)(\pi-1)\left(2b + \beta_2 + \beta_2^2 - b\beta_2^2 - b\beta_2 + 1\right) < 0$$

which holds if $\left(2b + \beta_2(1-b) + \beta_2^2(1-b) + 1\right) > 0$. The last condition is true. For $s = -1$, $(*)$ is expressed as

$$-\frac{4}{3}\frac{\pi}{c^4\alpha^2}(2c-3y)(\pi-1)\left(b\beta_2 - \beta_2 + \beta_2^2 - b\beta_2^2 + 1\right) < 0$$

which is holds if $1 + \beta_2((\beta_2 - 1)(1 - b) > 0$. This condition is true under our assumptions. Finally to show that positive expected profit obtains, it suffices to show that $\Omega^n(y/2)|_{s\in\{-1,1\}} = \pi y(2-\pi) > 0$. □

Proof of Proposition 5. The first order condition for c is given by $\Omega^{n\prime}(c) = -2\pi^2 - 4(1-\pi)\pi Ep^n + 2(1-\pi)\pi\frac{\partial Ep^n}{\partial c}(y-2c) = 0$. Consider π first. The proof follows that of Proposition 1 with Ep^n replacing p^n. We have that $\partial c/\partial\pi=$ sign $\left[-2\pi + (1-2\pi)\frac{\partial Ep^n(c)}{\partial c}(y-2c) - 2(1-2\pi)Ep^n(c)\right]$. Substituting from the first order condition, the latter expression can be written as

$$\begin{aligned}&\text{sign}[-2\pi + (1-2\pi)(\pi/(1-\pi) + 2Ep^n(c))\\&\quad - 2(1-2\pi)Ep^n(c)] = \text{sign}\left[-2 + \frac{1-2\pi}{1-\pi}\right] < 0.\end{aligned}$$

Thus, $\partial c/\partial\pi < 0$. Next, consider $\partial\Omega^n(c)/\partial\pi = 2(1-2\pi)y + 4\pi(y-c) + 2(1-2\pi)Ep^n(c)(y-2c)$. Clearly if $\pi \leqslant 5$, then $\partial\Omega^n(c)/\partial\pi > 0$. If $\pi > .5$, then $\partial\Omega^n(c)/\partial\pi \gtreqless 0$ as $0 \gtreqless -(y-2c\pi) + Ep^n(c)(y-2c)(2\pi-1)$. The

$$\begin{aligned}&\lim_{\pi\to 1} -(y-2c\pi) + Ep^n(c)(y-2c)(2\pi-1)\\&\quad = -(y-2c) + Ep^n(c)(y-2c) < 0,\ \text{and}\end{aligned}$$

$$\lim_{\pi\to .5} -(y-2c\pi) + Ep^n(c)(y-2c)(2\pi-1) = -(y-c) < 0.$$

Further,

$$\begin{aligned}&\frac{\partial}{\partial\pi}[-(y-2c\pi) + Ep^n(c)(y-2c)(2\pi-1)]\\&\quad = 2c + 2Ep^n(c)(y-2c) > 0.\end{aligned}$$

It follows that $\partial\Omega^n(c)/\partial\pi > 0$. Next consider α and note that

$$\begin{aligned}\frac{\partial Ep^n(c)}{\partial\alpha} &= -\frac{2}{c^2\alpha^3(s^2-4)}(b(1+s) + \beta_2(1-b)(s+\beta_2)+1) > 0 \text{ and}\\ \frac{\partial^2 Ep^n(c)}{\partial c\partial\alpha} &= \frac{4}{c^3\alpha^3(s^2-4)}(b(1+s) + \beta_2(1-b)(s+\beta_2)+1) < 0.\end{aligned}$$

Whence, $\frac{\partial\Omega^{n\prime}(c)}{\partial\alpha} = -4(1-\pi)\pi\frac{\partial Ep^n}{\partial\alpha} + 2(1-\pi)\pi\frac{\partial^2 Ep^n}{\partial c\partial\alpha}(y-2c) < 0$. Further,

$$\begin{aligned}\text{sign}\{\partial\Omega^n(c)/\partial\alpha\} &= \text{sign}\left\{2(1-\pi)\pi\left(\frac{\partial Ep^n(c)}{\partial\alpha}\right)(y-2c)\right\}\\ &= \text{sign}\left\{\frac{\partial Ep^n(c)}{\partial\alpha}\right\} > 0.\end{aligned}$$

For β_2, we have $\frac{\partial Ep^n(c)}{\partial\beta_2} = -\frac{1}{c^2\alpha^2(s^2-4)}(s+2\beta_2)(b-1) < 0$ and $\frac{\partial^2 Ep^n(c)}{\partial c\partial\beta_2} = \frac{2}{c^3\alpha^2(s^2-4)}(s+2\beta_2)(b-1) > 0$, so that $\frac{\partial\Omega^{n\prime}(c)}{\partial\beta_2} = -4(1-\pi)\pi\frac{\partial Ep^n}{\partial\beta_2} + 2(1-\pi)\pi\frac{\partial^2 Ep^n}{\partial c\partial\beta_2}(y-2c) > 0$. Moreover, sign $\{\partial\Omega^n(c)/\partial\beta_2\}$= sign $\left\{2(1-\pi)\pi\left(\frac{\partial Ep^n(c)}{\partial\beta_2}\right)(y-2c)\right\}$ = sign $\left\{\frac{\partial Ep^n(c)}{\partial\beta_2}\right\} < 0$. Finally, for b, we have $\frac{\partial Ep^n(c)}{\partial b} = -\frac{1}{c^2\alpha^2}\frac{\beta_2-1}{s^2-4}(s+\beta_2+1) > 0$ and $\frac{\partial^2 Ep^n(c)}{\partial c\partial b} = \frac{2}{c^3\alpha^2}\frac{\beta_2-1}{s^2-4}(s+\beta_2+1) < 0$, so that $\frac{\partial\Omega^{n\prime}(c)}{\partial b} = -4(1-\pi)\pi\frac{\partial Ep^n}{\partial b} + 2(1-\pi)\pi\frac{\partial^2 Ep^n}{\partial c\partial b}(y-2c) < 0$. The sign $\{\partial\Omega^n(c)/\partial b\}$= sign $\left\{2(1-\pi)\pi\left(\frac{\partial Ep^n(c)}{\partial b}\right)(y-2c)\right\}$ = sign $\left\{\frac{\partial Ep^n(c)}{\partial b}\right\} > 0$. □

Derivation of the comparative statics of Section 5.2

Given that $\Omega^i(c) = 2(1-\pi)\pi y + 2\pi^2(y-c) + 2(1-\pi)\pi(Ep^i)(y-2c), i = s,n$, and the fact that c is chosen optimally, the envelope theorem implies that $\partial\Omega^i(c)/\partial k = 2(1-\pi)\pi(y-2c)\partial Ep^i/\partial k, i = s,n, k = \Delta,\ell$. By $2(1-\pi)\pi(y-2c) > 0$, sign $\partial\Omega^i(c)/\partial k$= sign $\partial Ep^i/\partial k$.

The simultaneous solution yields ($s^2 = 1$)

$$\begin{aligned}T^n_{11} &= \frac{1}{3c}(2c - s + cs\Delta + cs\ell - 2),\\ L^n_{11} &= \frac{1}{3c}(2c\Delta - s + 2c\ell + cs - 2),\end{aligned}$$

$$\begin{aligned}T^n_{12} &= \frac{1}{3c}(2c - s\beta_2 + cs\Delta + cs\ell - 2),\\ L^n_{12} &= \frac{1}{3c}(2c\Delta - 2\beta_2 - s + 2c\ell + cs),\end{aligned}$$

$$\begin{aligned}T^n_{21} &= \frac{-1}{3c}(-2c + s + cs\Delta - cs\ell + 2),\\ L^n_{21} &= \frac{-1}{3c}(s + 2c\Delta - 2c\ell - cs + 2),\end{aligned}$$

$$\begin{aligned}T^n_{22} &= \frac{-1}{3c}(-2c + s\beta_2 + cs\Delta - cs\ell + 2),\\ L^n_{22} &= \frac{-1}{3c}(s + 2\beta_2 + 2c\Delta - 2c\ell - cs),\end{aligned}$$

$$\begin{aligned}&\frac{\partial}{\partial\Delta}(T^n_{11}) = \frac{s}{3}, \quad \frac{\partial}{\partial\Delta}(L^n_{11}) = \frac{2}{3},\\ &\frac{\partial}{\partial\Delta}(T^n_{12}) = \frac{s}{3}, \quad \frac{\partial}{\partial\Delta}(L^n_{12}) = \frac{2}{3},\end{aligned}$$

$$\begin{aligned}&\frac{\partial}{\partial\Delta}(T^n_{21}) = \frac{-s}{3}, \quad \frac{\partial}{\partial\Delta}(L^n_{21}) = \frac{-2}{3},\\ &\frac{\partial}{\partial\Delta}(T^n_{22}) = \frac{-s}{3}, \quad \frac{\partial}{\partial\Delta}(L^n_{22}) = \frac{-2}{3},\end{aligned}$$

$$\begin{aligned}&\frac{\partial}{\partial\ell}(T^n_{11}) = \frac{s}{3}, \quad \frac{\partial}{\partial\ell}(L^n_{11}) = \frac{2}{3},\\ &\frac{\partial}{\partial\ell}(T^n_{12}) = \frac{s}{3}, \quad \frac{\partial}{\partial\ell}(L^n_{12}) = \frac{2}{3}\end{aligned}$$

$$\begin{aligned}&\frac{\partial}{\partial\ell}(T^n_{21}) = \frac{s}{3}, \quad \frac{\partial}{\partial\Delta}(L^n_{21}) = \frac{2}{3},\\ &\frac{\partial}{\partial\Delta}(T^n_{22}) = \frac{s}{3}, \quad \frac{\partial}{\partial\Delta}(L^n_{22}) = \frac{2}{3},\end{aligned}$$

Using these definitions and the incentive compatibility conditions,

$$\begin{aligned}\frac{\partial Ep^n}{\partial\Delta} = .5&\left[\frac{b}{c^n}\left(\frac{s}{3} + \frac{2}{3}\right) + \frac{(1-b)\beta_2}{c^n}\left(\frac{s}{3} + \frac{2}{3}\right)\right]\\ &+ .5\left[\frac{b}{c^n}\left(\frac{-s}{3} + \frac{-2}{3}\right) + \frac{(1-b)\beta_2}{c^n}\left(\frac{-s}{3} + \frac{-2}{3}\right)\right]\\ &+ .5\left[b\left(\frac{4\Delta}{3}\right) + (1-b)\left(\frac{4\Delta}{3}\right)\right] = \frac{2\Delta}{3}.\end{aligned}$$

Next consider

$$\frac{\partial Ep^n}{\partial \ell} = .5\left[\frac{b}{c^n}\left(\frac{s}{3}+\frac{2}{3}\right)+\frac{(1-b)\beta_2}{c^n}\left(\frac{s}{3}+\frac{2}{3}\right)\right] + .5\left[\frac{b}{c^n}\left(\frac{s}{3}+\frac{2}{3}\right)+\frac{(1-b)\beta_2}{c^n}\left(\frac{s}{3}+\frac{2}{3}\right)\right] + 0.5\left[\left(b\left(L^n_{11}+L^n_{21}\right)+(1-b)\left(L^n_{12}+L^n_{22}\right)\right]\right.$$
$$= \left[b+(1-b)\beta\right]\left(\frac{s+2}{3c^n}\right)+0.5\left[\left(b\left(L^n_{11}+L^n_{21}\right)+(1-b)\left(L^n_{12}+L^n_{22}\right)\right]\right..$$

In the sequential mode we have

$$T^s = \frac{1}{3c(c+1)}(2.0c+2.0c^2+cs\Delta+cs\ell-1.0c^2s\Delta+c^2s\ell-4.0),$$

$$L^s_{11} = \frac{1}{6c(c+1)}(4.0c\Delta-4.0s-4.0c+4.0c\ell+cs^2+2.0c^2s+4.0c^2\Delta+4.0c^2\ell+2.0cs+s^2-2.0c^2s^2\Delta-4.0),$$

$$L^s_{21} = \frac{1}{6c(c+1)}(4.0c\ell-4.0s-4.0c\Delta-4.0c+c+2.0c^2s-4.0c^2\Delta+4.0c^2\ell+2.0cs+1+2c\Delta-4.0).$$

Note that for $L^s_{21} > 0$, we require at $\ell = 1$

$$(4.0c-4.0s-4.0c\Delta-4.0c+c+2.0c^2s-4.0c^2\Delta+4.0c^2+2.0cs+1+2c\Delta-4.0) > 0.$$

If $s = -1$, this implies

$$c^2(2-4\Delta)-c(1+2\Delta)+1>0.$$

For this to be positive for $c > 1$, we need to keep $\Delta \in (0,1)$ small. If $c > 1$, then $\Delta > 0.5$ implies that $c^2(2-4\Delta) < 0$, $-c(1+2\Delta) < 0$ with $|c(1+2\Delta)| > 1$. Thus $L^s_{21} < 0$ and a contradiction. When $c > 1$, it must be that $\Delta < 0.5$ for $L^s_{21} > 0$.

We have that

$$\frac{\partial T^s}{\partial \Delta} = \frac{s(c-1)}{-3(c+1)}, \frac{\partial T^s}{\partial \ell} = \frac{s}{3},$$
$$\frac{\partial}{\partial \Delta}\left(L^s_{11}\right) = \frac{c+2}{3(c+1)}, \frac{\partial}{\partial \ell}\left(L^s_{11}\right) = 2/3,$$
$$\frac{\partial}{\partial \Delta}\left(L^s_{21}\right) = \frac{-(2c+1)}{3(c+1)}, \frac{\partial}{\partial \ell}\left(L^s_{21}\right) = 2/3.$$

Computing

$$\frac{\partial Ep^s}{\partial \Delta} = \frac{-s(c^s-1)}{3c^s(c^s+1)} + 0.5(\Delta).$$

Next,

$$\frac{\partial Ep^s}{\partial \ell} = \frac{s}{3c^s}+0.5\left(L^s_{11}+L^s_{21}\right), \quad \text{where } L_{11}+L_{21}$$
$$= \frac{1}{3c^s(c^s+1)}4.0c^s\ell-4.0s-4.0c^s+c^ss^2+2.0c^{s2}s+4.0c^{s2}\ell+2.0c^ss+s^2-1.0c^{s2}s^2\Delta+c^ss^2\Delta-4.0.$$

Evaluating at $\ell = 1$ and simplifying we obtain

$$\frac{\partial Ep^s}{\partial \ell} = \frac{1}{6c^s(c^s+1)}(c^s\Delta-2.0s-3.0c^s+4.0c^s+2.0c^{s2}s-1.0c^{s2}\Delta+4.0c^{s2}+4.0c^ss-3.0)$$

and if $s = -1$, we have

$$\frac{\partial Ep^s}{\partial \ell} = -\frac{1}{6c^s(c^s+1)}(3.0c^s-1.0c^s\Delta+c^{s2}\Delta-2.0c^{s2}+1.0).$$

References

Bulow, Jeremy, Geanakoplos, John D., Klemperer, Paul D., 1985. Multimarket oligopoly: strategic substitutes and complements. Journal of Political Economy 93, 488–510.
Dessein, Wouter, Santos, Tano, 2006. Adaptive organizations. Journal of Political Economy 114, 956–995.
Mathias, Dewatripont, Jean, Tirole, 2005. Modes of communication. Journal of Political Economy 113, 1217–1238.
Dryer, Jeffery H., Nobeoka, Kentaro, 2000. Creating and managing a high-performance knowledge-sharing network: the Toyota case. Strategic Management Journal 21, 345–367.
Itoh, Hideshi, 1991. Incentives to help in multi-agent situations. Econometrica 59, 611–636.
Marschak, Jacob, Radner, Roy, 1972. Economic Theory of Teams. Yale University Press, New Haven, CT.
Thurm, Scott, 2006. Companies struggle to pass on knowledge that workers acquire. Wall Street Journal, Monday, January 23, 2006.
Van Zandt, Timothy, 1999. Decentralized information processing in the theory of organizations. In: Contemporary Economic Issues, vol. 4: Economic Behavior and Design, edited by Murat Sertel: Macmillan Press, pp. 125–160.