INTER-TEMPORAL AGGREGATION AND INCENTIVES

Anil Arya
The Ohio State University

Jonathan Glover
Carnegie Mellon University

Pierre Jinghong Liang
Carnegie Mellon University

August 2003

Abstract

Inter-temporal aggregation results in a summarization of information and a natural delay in the release of information. We study a principal-agent model and show that inter-temporal aggregation can be an optimal feature of a performance evaluation system. We also contrast the value of additional information when existing information is inter-temporally aggregated with its value when existing information is obtained period-by-period. We observe that valuing additional information is subtle. In particular, it is possible the principal will find that an additional independent signal is valueless under inter-temporal aggregation but valuable if existing information is obtained periodically. That is, additional information can be more helpful when other information sources abound than when they are scarce.

1. Introduction

Accounting information exhibits inter-temporal aggregation in that it is associated with both summary and delay. The accountant's mantra of double-entry accrual bookkeeping conveys these aspects: linear journal entries summarize, and carefully crafted accruals time information release. There have been calls for "real-time" detailed reporting, which leaves summarization to individuals users. It seems the only thing standing in the way of such continuous reporting is the cost of preparing and disseminating information, which, with advances in information technology, is less of an issue. In contrast, this paper emphasizes that a move away from inter-temporal aggregation towards timeliness and detail can be associated with an increase in incentive costs.

To study inter-temporal aggregation and incentives, we model a two-period moral hazard setting wherein either period-by-period output is measured (a "frequent" evaluation regime) or an aggregate output over the two periods is measured (an "infrequent" evaluation regime).[1] Aggregation under infrequent evaluation is associated with a cost: in compensating the agent, the principal is restricted to using a summarized proxy. However, the delay under infrequent evaluation is associated with a benefit: when choosing his second-period act, the agent's information (and, hence his opportunism) is limited because of his ignorance about the first period's output.[2]

We employ a model in which it might seem there are no incentive benefits to delay. There are no productive interdependencies across periods, and the participants' preferences ensure there is no demand for consumption smoothing. Nevertheless, the incentive benefit

---

1 Feltham (1972) studies a single-person decision problem in which infrequent reporting translates into less-informed decision making. In this paper, we revisit the issue of periodicity of reporting in a strategic setting.

2 Banker and Datar (1989) use a single-period agency model (as in Holmstrom, 1979) to study the cost of aggregation. Banker and Datar identify necessary and sufficient conditions on the joint density function of signals under which linear aggregation is optimal. Further, the linear weights are assigned an intuitive interpretation in terms of signal sensitivity and precision.

persists. Our proposition and examples illustrate the delay benefit and summarization cost associated with inter-temporal aggregation.

From a technical perspective, the benefit of delayed information release is fairly obvious—there are simply less incentive compatibility constraints to satisfy. That this benefit would have a strictly positive value is less obvious. The underlying force is as follows. Without delay, the second-period bonus can be used to motivate only second-period effort, since second-period incentives have to hold for all first-period outcome realizations. With delay, the second-period bonus can be paid for outcomes that are indicative of both first- and second-period effort, allowing second-period incentives to spillover (spillback) into the first period. This reduces the required bonus needed for a good first-period outcome.

Another way to view the result is that releasing information late increases the uncertainty at the time the agent takes his second-period act, thereby creating more opportunities for risk sharing. Roughly stated, delay increases the sample size that can be used to infer the agent's supply of inputs. The principal's inference exercise is not disabled by independence across periods in much the same manner as an independent sample does not hinder, but rather helps, a statistician gauge population parameters. Likewise, securities with independent returns do not prevent investors from obtaining diversification gains. Our infrequent evaluation regime yields similar risk-sharing gains but only at the expense of summarized contractual data.

Gigler and Hemmer (1998) have made the observation that delaying but fully preserving public information can be optimal if the agency relationship is repeated enough times. In contrast, Gigler and Hemmer (2002) study a two-period model in which aggregation across divisions can be valuable as a means of reducing the information available to agents when choosing their second-period supply of productive inputs. Our focus on inter-temporal aggregation borrows elements from both these papers. (See Gigler and Hemmer (2002, pp. 71-72) for a discussion of this and other differences between our work and theirs.)

Aggregation is a pervasive theme in accounting (see, for example, Demski 1994; Ijiri 1971, 1975; Sunder 1997). While it is presumably easier to appreciate the cost of aggregation, there are also potential gains to aggregation. One benefit to aggregation has to do with bounded rationality: computational costs and information overload can make detailed information less valuable than aggregated information. A second potential benefit to aggregation arises when individual items are measured with errors. Aggregation sometimes allows errors in the individual items to cancel out (Datar and Gupta, 1994; Grunfeld and Griliches, 1960; Lim and Sunder, 1991). A third reason is that the aggregation process itself may add information (Sunder, 1997). Finally, agency models have studied settings in which the principal's limited ability to commit leads to coarse information being optimal (see, for example, Arya, Glover, and Sivaramakrishnan 1997; Cremer 1995; and Sappington 1986).

We also compare the value of additional information when existing information is inter-temporally aggregated with its value when existing information is obtained period-by-period. We observe that valuing additional information is subtle. In particular, it is possible the principal will find that an additional independent signal is valueless under inter-temporal aggregation but valuable if existing information is obtained periodically.[3] That is, additional information can be more helpful when other information sources abound than when they are scarce.

The remainder of the paper is organized into three sections. Section 2 studies the optimal periodicity of performance measurement. Section 3 studies the value of additional information. Section 4 concludes the paper.

---

[3] In the case of dependent signals, the notion of informativeness/conditional controllability developed in the standard moral hazard model (e.g., Holmstrom, 1979) can be used to derive a similar result.

2. Periodicity of Performance Measurement

2.1 Model

The model is similar to Fellingham, Newman, and Suh (1985) and the benchmark setting in Demski (1998) in that it employs CARA (constant absolute risk aversion) preferences and an iid production technology. A risk-neutral principal contracts with a risk- and effort-averse agent for two periods. The agent chooses an action a_t, $a_t \in \{L,H\}$, in period t, t = 1,2. The agent's action choice is subject to moral hazard; the principal does not observe either period's action. The publicly observable output in each period is denoted by x_t, $x_t \in \{0,1,...,n\}$.[4]

Each period's output is a function of that period's act and a stochastic state variable θ. The first-period act does not affect the second period's output, nor is the first-period output informative of the second-period output. Hence, $Pr(x_1,x_2|a_1,a_2) = Pr(x_1|a_1)Pr(x_2|a_2)$. In addition to this independence assumption, we assume the distributions are identical across periods: $Pr(x_1=m|a_1=i) = Pr(x_2=m|a_2=i)$. Also, the distributions satisfy first-order stochastic dominance in the agent's act.

The principal and the agent either both observe output frequently or infrequently. A frequent evaluation regime $R(x_1,x_2)$ is one in which performance is measured at the end of each period. An infrequent evaluation regime $R(\phi,x_1+x_2)$ is one in which no information (ϕ) is obtained at the end of period one and x_1+x_2 is obtained at the end of the second period. (Because infrequent evaluation involves both aggregation and delay, we use the terms infrequent evaluation and inter-temporal aggregation interchangeably.) Hence, the contract between the principal and the agent depends on x_1 and x_2 in the $R(x_1,x_2)$ regime and only on x_1+x_2 in the $R(\phi,x_1+x_2)$ regime.

---

4 To be fair, we derive our proposition for the binary outcome case, which is special in that there are no costs to aggregation. We then use trinary outcome examples to illustrate both the benefits and the costs of inter-temporal aggregation.

Denote by s(·) the contractual payment the principal makes to the agent. The agent's preferences over compensation and effort are represented by a von Neumann-Morgenstern utility function $U[s-a_1-a_2] = -e^{-c(s-a_1-a_2)}$, where $c > 0$ is the risk-aversion parameter. Note a_1 and a_2 are stated in monetary equivalents, with L < H. We assume the agent cares about the total payment he receives at the end of two periods, not how payments are split across periods. Hence, without loss of generality, assume s(·) is paid to the agent at the end of period two.

There are several advantages to using constant absolute risk aversion (CARA) to model the agent's preferences. First, CARA focuses on the first-order effect of risk aversion by eliminating the second-order effect of changes in risk aversion due to wealth. Second, it ensures contracts involving ex ante randomization cannot be optimal (Arya, Fellingham, and Young, 1993). Third, with this representation, preference-driven demands for a decision maker to condition decisions on past outcomes and for a principal-agent relationship to make use of long-term contracts are eliminated (Amershi, Demski, and Fellingham, 1985; Fellingham, Newman, and Suh, 1985).

The principal's preferences are given by $P \cdot (x_1+x_2) - s$. P is the per unit value of output to the principal. We assume P is sufficiently large that motivating H in both periods is optimal under either evaluation regime. Also, we assume $Pr(x_t|a_t=H) > 0$ for all x_t. These assumptions ensure there is a non-trivial incentive problem.

The principal's objective is to minimize the expected cost of motivating the manager to choose H in each period. The minimization problem is constrained in that it must be in the agent's own best interest to join the firm and choose H in each period rather than not join the firm (the individual rationality constraint (IR)) or join the firm and choose some other act combination (the family of incentive compatibility constraints (IC)). Denote by M the agent's two-period reservation certainty equivalent. The minimization programs under $R(x_1,x_2)$ and $R(\phi,x_1+x_2)$ are as follows.

$R(x_1,x_2)$:

$$\min_s \sum_{x_1}\sum_{x_2} \Pr(x_1|H)\Pr(x_2|H)s(x_1,x_2)$$

$$\sum_{x_1}\sum_{x_2} \Pr(x_1|H)\Pr(x_2|H)U[s(x_1,x_2) - H - H] \geq U[M] \qquad \text{(IR)}$$

$$\Pr(0|H)\sum_{x_2} \Pr(x_2|H)U[s(0,x_2) - H - H] + \Pr(1|H)\sum_{x_2} \Pr(x_2|H)U[s(1,x_2) - H - H] \geq$$
$$\Pr(0|i)\sum_{x_2} \Pr(x_2|j)U[s(0,x_2) - i - j] + \Pr(1|i)\sum_{x_2} \Pr(x_2|k)U[s(1,x_2) - i - k]$$
$$i,j,k = L,H \qquad \text{(IC)}$$

$R(\phi,x_1+x_2)$:

$$\min_s \sum_{x_1}\sum_{x_2} \Pr(x_1|H)\Pr(x_2|H)s(x_1+x_2)$$

$$\sum_{x_1}\sum_{x_2} \Pr(x_1|H)\Pr(x_2|H)U[s(x_1+x_2) - H - H] \geq U[M] \qquad \text{(IR)}$$

$$\sum_{x_1}\sum_{x_2} \Pr(x_1|H)\Pr(x_2|H)U[s(x_1+x_2) - H - H] \geq$$
$$\sum_{x_1}\sum_{x_2} \Pr(x_1|i)\Pr(x_2|j)U[s(x_1+x_2) - i - j] \qquad i,j = L,H \qquad \text{(IC)}$$

To disentangle the effects of aggregation and delay, it is helpful to compare $R(\phi,x_1+x_2)$ to $R(x_1,x_2)$ in two steps. First, compare $R(\phi,x_1+x_2)$ to $R(\phi,(x_1,x_2))$. Here, the only difference is aggregation. It is well-known that loss of information can be costly in a standard moral hazard setting. The $R(\phi,x_1+x_2)$ program reflects this by effectively adding constraints requiring $s(x_m,x_n) = s(x_n,x_m)$ to the $R(x_1,x_2)$ program.

Second, compare $R(\phi,(x_1,x_2))$ to $R(x_1,x_2)$. Here, the only difference is timing. Delay can be beneficial because it reduces the agent's opportunism options. The $R(\phi,x_1+x_2)$ program reflects this by deleting incentive constraints from the $R(x_1,x_2)$ program in which $j \neq k$. In the next section, the principal chooses between frequent and infrequent evaluation by comparing the cost of aggregation with the benefit of delay.

2.2 Result

In the binary CARA setup, there is no cost of aggregation. This is easy to see. The optimal contract in the $R(x_1,x_2)$ regime is a two-fold repetition of the optimal single-period contract (Fellingham, Newman, and Suh, 1985). Denote the single-period solution (with a single-period reservation certainty equivalent of M/2) by $s^*(0)$ and $s^*(1)$. Hence, $s(0,0) = 2s^*(0)$, $s(0,1) = s(1,0) = s^*(0) + s^*(1)$, and $s(1,1) = 2s^*(1)$.

Note the events (0,1) and (1,0) are pooled under aggregation but, since the payments in these events are identical, pooling is costless. Thus, the optimal contract under $R(x_1,x_2)$ is feasible in the $R(\phi,x_1+x_2)$ regime. The question is: is this contract optimal in the $R(\phi,x_1+x_2)$ regime? The following proposition says no. (The proof is provided in the Appendix.)

Proposition. For the case of binary outputs in each period, the principal is strictly better off when performance is measured at the end of two periods (infrequent evaluation) instead of at the end of each period (frequent evaluation).

The optimal contract in the $R(\phi,x_1+x_2)$ regime exhibits "decreasing returns to good news" (Demski, 1998, Fact 2), while the Fellingham, Newman, and Suh (1985) contract exhibits constant returns to scale. In other words, when the optimal $R(x_1,x_2)$ contract is employed in the $R(\phi,x_1+x_2)$ regime, any (i,j) act combination provides the agent with his reservation utility, while, under the optimal $R(\phi,x_1+x_2)$ contract, the (LL) constraint provides the agent with strictly less than his reservation utility.[5] (The notation (i,j) is used to indicate (a_1,a_2), i,j = L,H.)

---

5 For an aggregation setting, Christensen, Demski, and Frimor (2002) introduce a Concavity of Aggregate Distribution Function Condition (CADFC). The CADFC is the multi-outcome version of Grossman and Hart's (1983) Concavity of Distribution Function Condition. This condition is sufficient but not necessary for only adjacent incentive compatibility constraints to bind. In our CARA binary outcome setting, even though this condition is not always satisfied, only adjacent incentive constraints come into play.

The phrase "decreasing returns to good news" means the increase in the certainty equivalent of payments when the agent's acts changes from (L,H) to (H,H) is less than when the acts change from (L,L) to (L,H). The same is true when (L,H) is replaced with (H,L). This follows from: (i) the probability distributions over aggregated output satisfying FOSD and (ii) the dollar payments also exhibiting decreasing returns to good news.

The result can also be explained in terms of diversification. The principal pays the agent a risk premium to compensate the agent for the noisiness of the performance measure, i.e., for the stochastic portion of the performance measure. When performance measures are aggregated, so are the stochastic terms. Further, delay implies both stochastic terms are unknown to the agent when choosing either of his acts. The risk associated with the sum of random variables can be lower than the sum of the risk associated with each individual random variable.[6] Such "diversification" reduces the risk premium required in the $R(\phi,x_1+x_2)$ regime relative to the $R(x_1,x_2)$ regime.

While, in the binary-outcome setting, there is only a benefit and no cost to inter-temporal aggregation, with more outcomes there are both costs and benefits.[7] However, inter-temporal aggregation can still be strictly preferred. Consider the following three outcome ($x_t \in \{0,1,2\}$) example.

Example 1

Parameters

L = 5000; H = 10000; M = 20000; c = 0.0001
Pr(0|L) = 0.36, Pr(1|L) = 0.48, Pr(2|L) = 0.16
Pr(0|H) = 0.04, Pr(1|H) = 0.32, Pr(2|H) = 0.64

---

6 Here, it helps to think of portfolio diversification and to recall that diversification comes into play for all correlation values less than 1 (including 0 correlation).

7 Alternatively, one can maintain the binary setting and change the agent's preferences to make this point. In particular, if the agent's preferences are not domain additive, the agent will care not only about the total payments but also about how payments are split across periods, i.e., a demand for consumption smoothing is introduced.

Solution

$R(x_1,x_2)$ regime
$s(m,n) = s^*(m) + s^*(n)$, where $s^*(0) = 9463$, $s^*(1) = 20131$, $s^*(2) = 21168$
Expected payment = 40736
$R(\phi,x_1+x_2)$ regime
s(0) = 3801, s(1) = 33522, s(2) = 38742, s(3) = 40975, s(4) = 41699
Expected payment = 40678

In Example 1, aggregation is costly because the principal cannot distinguish the events (0,2) and (2,0) from (1,1). As mentioned earlier, a natural way to isolate this cost is by comparing the $R(\phi,(x_1,x_2))$ regime with the $R(\phi,(x_1+x_2))$ regime.[8] Under the $R(\phi,(x_1,x_2))$ regime, the solution is s(0,0) = 5769, s(0,1) = s(1,0) = 33577, s(0,2) = s(2,0) = 35030, s(1,1) = 40194, s(1,2) = s(2,1) = 40970, and s(2,2) = 41690. The expected payment is 40635. Hence, the cost of aggregation is 40678 - 40635 = 43.

The benefit of delay can be isolated by comparing the $R(\phi,(x_1,x_2))$ regime with the $R(x_1,x_2)$ regime. In the former case, there are fewer incentive compatibility constraints—the agent has less information on which to base a_2. The benefit is 40736 - 40635 = 101. Since the benefit of delay exceeds the cost of aggregation, infrequent evaluation is strictly preferred. Of course, the tradeoff can go the other way. The cost of aggregation may exceed the delay benefit, making frequent evaluation optimal. Such an example is presented next.

Example 2

Parameters

L = 5000; H = 10000; M = 20000; c = 0.0001
Pr(0|L) = 0.49, Pr(1|L) = 0.42, Pr(2|L) = 0.09
Pr(0|H) = 0.04, Pr(1|H) = 0.32, Pr(2|H) = 0.64

8 In Example 1 (and Example 2), under both $R(\phi,(x_1,x_2))$ and $R(\phi,(x_1+x_2))$, only the adjacent incentive constraints (L,H) and (H,L) bind. In contrast, under $R(x_1,x_2)$, all incentive constraints bind.

Solution

$R(x_1,x_2)$ regime
$s(m,n) = s^*(m) + s^*(n)$, where $s^*(0) = 11442$, $s^*(1) = 20181$, $s^*(2) = 20786$
Expected payment = 40438
$R(\phi,x_1+x_2)$ regime
s(0) = 7859, s(1) = 33750, s(2) = 38757, s(3) = 40802, s(4) = 41280
Expected payment = 40450

A final note regarding the proposition: while agent risk aversion is used to introduce contracting friction in our model, the proposition also holds under agent risk neutrality with nonnegativity constraints on the payments. In fact, absent the need for risk-sharing, there are only two distinct payments irrespective of the number of outcomes: a bonus payment for the "best news" (highest likelihood ratio) event and 0 for all other outcomes. This has the following implications. First, the binary-outcome assumption can be readily dispensed with when writing an analogous proposition, since from an incentive perspective there are, in effect, only two events. Second, since the payments no longer exhibit decreasing returns to good news, neither do the expected payments. In fact, the complete loading of payments on the most favorable event (i.e., drastic response to best news) translates into only the non-adjacent (L,L) constraint binding in the risk-neutral setting. The following example illustrates these points.

Suppose $a_L = 0$, $a_H = 1$, Pr(2|H) = 0.75, Pr(1|H) = 0.15, Pr(0|H) = 0.10, Pr(2|L) = 0.25, Pr(1|L) = 0.50, and Pr(0|L) = 0.25. In the one-shot version of this example, the required contract is s(2) = 1/(0.75 - 0.25) = 2 and $s(\cdot) = 0$ otherwise, yielding an expected payment of 0.75(2) = 1.5. Under frequent evaluation, performance measurement is repeated and so is the optimal contract —the expected payment is 3. Under infrequent evaluation, s(4) = 2/[(0.75)(0.75) - (0.25)(0.25)] = 4 and $s(\cdot) = 0$ otherwise, yielding a lower expected payment of (0.75)(0.75)4 = 2.25.

3. Value of Additional Information

In this section, we study the impact of inter-temporal aggregation on the value of additional information. The issue seems relevant, since accounting both routinely aggregates over time is often used in conjunction with other information (e.g., stock prices).

In the binary outcome CARA setting, we know the principal strictly prefers the $R(\phi,x_1+x_2)$ regime to the $R(x_1,x_2)$ regime. Now introduce an additional piece of information, say signal y, which is publicly available at the end of period one. Two questions arise. First, is this ranking of information systems maintained, i.e., is delayed aggregated information strictly preferred to periodic information? Second, in which regime is y more valuable?

It is tempting to pursue the following line of reasoning. In both regimes, y has the same effect on the agent's incentive compatibility constraints for choosing a_2. This implies the ranking of regimes is unchanged. Regarding the value of the additional signal, with (x_1,x_2) the principal has more information than with x_1+x_2. Hence, y is more helpful in the $R(\phi,x_1+x_2)$ regime.

The flaw in the above logic is highlighted by considering a crude example in which y is simply x_1. Clearly, the expected payments are the same under $R(y,x_1+x_2)$, $R((y,x_1),x_2)$, and $R(x_1,x_2)$.[9] That is, with $y = x_1$, the principal is indifferent between infrequent and frequent measurement of x. Also, note the value of y is negative in the $R(y,x_1+x_2)$ regime and 0 in the $R((y,x_1),x_2)$ regime; y is more valuable (less costly) when x is measured frequently. (The value of y is the difference in expected payments with and without y.)

Consider another crude example in which y is a perfect monitor of a_1 ($y = a_1$). This implies there is no moral hazard problem regarding a_1. Under $R((y,x_1),x_2)$, the problem decomposes—the first-best solution is obtained in the first period and the single-period, second-best solution is obtained in the second period. Under $R(y,x_1+x_2)$, aggregation simply

---

9 With y, the $R(\phi,x_1+x_2)$ and $R(x_1,x_2)$ regimes are labeled $R(y,x_1+x_2)$ and $R((y,x_1),x_2)$, respectively.

adds noise to the second-period performance measure. That is, in the presence of y, the principal prefers to measure x period-by-period. Moreover, under such measurement, y is more valuable.

Finally, consider a less extreme example in which y depends on the agent's first-period act but is otherwise independent of x_1 and x_2. Here, y is informative about a_1 (conditionally controllable), given either (x_1,x_2) or (x_1+x_2). In Example 3, (i) the presence of y reverses the ranking over regimes, and (ii) y is valueless when output is aggregated at the end of two periods but valuable when output is measured periodically.

Example 3

Parameters

L = 5000; H = 10000; M = 20000; c = 0.0001
$Pr(x_t = 1|a_t = L) = 0.20$, $Pr(x_t = 1|a_t = H) = 0.80$
$Pr(y = 1|a_1 = L) = 0.40$, $Pr(y = 1|a_1 = H) = 0.60$

Solution

$R(x_1,x_2)$ regime

$s(x_1,x_2)$: s(0,0) = 27535.19, s(0,1) = s(1,0) = 36204.125,
s(1,1) = 44873.06
Expected payment = 41405.485

$R((y,x_1),x_2)$ regime

$s(y,x_1,x_2)$: s(0,0,0) = 25386.075, s(0,0,1) = 34055.010,
s(0,1,0) = 35304.402, s(0,1,1) = 43973.337
s(1,0,0) = 32343.648, s(1,0,1) = 41012.583,
s(1,1,0) = 35534.610, s(1,1,1) = 44203.545
Expected payment = 41201.293

$R(\phi,x_1+x_2)$ and $R(y,x_1+x_2)$ regimes

$s(y,x_1+x_2) = s(x_1+x_2)$ (It is optimal to ignore y.)
$s(x_1+x_2)$: s(0) = 24798.90, s(1) = 38277.12, s(2) = 43817.08
Expected payment = 41283.565

Value of y with (x_1,x_2) = 204.192
Value of y with (ϕ,x_1+x_2) = 0

The use of additional information can be subtle in the presence of inter-temporal aggregation. Under $R(\phi, x_1+x_2)$, the (IC) constraints involving act combinations (L,H) and (H,L) bind while (L,L) does not bind.[10] This implies the payment schedule that ensures the agent provides high effort in the second period also ensures high effort in the first period. In other words, the first-period high effort can be acquired "for free," so to speak. It is then not surprising additional information about a_1 is valueless.[11]

Generally, the arrival of a signal mid-game has two effects on contracting. First, the principal has additional information about the agent's first-period action (a positive effect). Second, more (IC) constraints regarding the agent's second-period action come into play (a weakly negative effect). In Example 3, the negative effect outweighs the positive when x is inter-temporally aggregated, so the optimal contract chooses to ignore y. Put differently, y helps with the first-period control problem because it provides another source of information about a_1. On the other hand, y exacerbates the second-period control problem because it reduces opportunities for risk sharing. The result of this tradeoff depends on whether or not x is inter-temporally aggregated. When x is not inter-temporally aggregated, the benefit of using y outweighs the cost. Even though y is conditionally independent of x_1 and x_2, there is an interaction between the optimal use of additional information and the level of inter-temporal aggregation.

To summarize, in Example 3, when output is aggregated, an informative signal that is produced mid-game is ignored. If the informative signal were produced at the end of the second period, it would always be used a la Holmstrom (1979). This is the feature of the

---

10 Under $R(y,(x_1+x_2))$, the number of (IC) constraints expand and are, for example, of the form (a_1=H; if y=1,a_2=H; if y=0,a_2=L), which we label as (H, HL). Under $R(y,(x_1+x_2))$ in Example 3, the binding constraints are (H, HL), (H,LH), (H,LL), and (L,HH). Under $R(x_1,x_2)$, all (IC) constraints bind while, under $R((y,x_1),x_2)$, the only nonbinding (IC) constraint is the one that ensures the agent chooses $a_2 = a_H$ when y = 0 and x_1 = 0.

11 A similar point has been made in the literature, although in other contexts. In a two-period setting without a repeated technology, Liang (2000) shows the value of additional information depends on which period's moral hazard is more critical (p. 446). In a single-period-multiple-task setting, Demski (1994) shows certain (IC) constraints may "free ride" other (IC) constraints, which can result in "bad measures driving out good measures" (Chapter 22, p. 577).

example we find intriguing.

4. Conclusion

This paper studies the incentive benefit to aggregation. In our model, the cost of inter-temporal aggregation is due to summarized (less) post-decision information that can be used in controlling the agent. The benefit is due to delayed information release that lessens the agent's pre-decision information at the time of choosing his act—this makes the control problem less severe. We rigged our model to eliminate any obvious interdependencies in the control problems across periods. CARA preferences and uncorrelated production were chosen with this in mind.

We also studied the value of additional information, taking disaggregated or aggregated performance evaluation as given. We observe that the optimal use of additional information interacts with how existing information is aggregated, even when the additional information is uncorrelated with the existing information. In particular, inter-temporal aggregation of an existing signal can make it optimal not to use an additional informative signal that is observed by both the principal and the agent at the end of the first period.

One benefit to aggregation we think would be interesting to model in future work is that the aggregation process (the choice of a particular aggregation rule) can be used to add information.

> The free, universal availability of an aggregation function is not typical in accounting contexts....To illustrate, in the process of aggregation, accountants add their knowledge and judgment about similarities and dissimilarities of various accounts. To write a single line in the aggregated balance sheet (say, inventories—$1 million), it is not sufficient to have a detailed list of each of the firm’s resources. One must also know the characteristics of all the resources and be able to decide which items from the detailed list could be usefully aggregated into a single line item labeled inventory on the balance sheet. This task cannot be performed by a lay person. An accounting expert brings special knowledge of aggregation functions and adds this information to the disaggregated data in order to arrive at the aggregated balance sheet. Sunder (1997, p. 89).

Appendix

Proof of Proposition.

Let $R'(\phi,x_1+x_2)$ denote a restricted version of program $R(\phi,x_1+x_2)$ which differs from $R(\phi,x_1+x_2)$ in that the constraint $s(1) - s(0) \leq s(2) - s(1)$ is added. The proof then proceeds in two steps. In step 1, we show the principal's expected compensation cost under $R(\phi,x_1+x_2)$ is strictly less than under $R'(\phi,x_1+x_2)$. In step 2, we show the principal's expected compensation cost under $R(x_1,x_2)$ is at least as high as under $R'(\phi,x_1+x_2)$. The two steps together imply the proposition.

Step 1. By construction, expected compensation cost under $R(\phi,x_1+x_2)$ is weakly lower than under $R'(\phi,x_1+x_2)$. The strict ordering follows from the fact that the unique optimal solution to $R(\phi,x_1+x_2)$ prescribes $s(1) - s(0) > s(2) - s(1)$ and is, hence, not feasible in the more constrained $R'(\phi,x_1+x_2)$. See Fact 2 and Footnote 9 in Demski (1998) for a proof that $s(1) - s(0) > s(2) - s(1)$.

Step 2. From Fellingham et al. (1985), the optimal contract in the $R(x_1,x_2)$ regime can be characterized as a two-fold repetition of the optimal single-period contract, i.e., $s(m,n) = s^*(m) + s^*(n)$. This solution is also feasible under $R'(\phi,x_1+x_2)$ because: (i) there are only three distinct payments, and the events corresponding to these payments are fully preserved under $R'(\phi,x_1+x_2)$ and (ii) $s(1) - s(0) = s(2) - s(1) = s^*(1) - s^*(0)$ and, hence, the constraint ordering the difference in payments is satisfied. ■

References

AMERSHI, A., J. DEMSKI, AND J. FELLINGHAM. "Sequential Bayesian Analysis in Accounting Settings," *Contemporary Accounting Research* 1, 1985, pp. 176-192.

ARYA, A., J. FELLINGHAM, AND R. YOUNG. "Preference Representation and Randomization in Principal Agent Contracts," *Economic Letters* 42, 1993, pp. 25-30.

ARYA, A., J. GLOVER, AND K. SIVARAMAKRISHNAN. "The Interaction between Decision and Control Problems and the Value of Information," *The Accounting Review*, Fall 1997, pp. 561-574.

BANKER, R. AND S. DATAR. "Sensitivity, Precision, and Linear Aggregation of Signals for Performance Evaluation," *Journal of Accounting Research* 27, 1989, pp. 21-39.

CREMER, J., "Arm's Length Relationships," *Quarterly Journal of Economics*, 1995, pp. 275-295.

DATAR, S. AND M. GUPTA. "Aggregation, Specification and Measurement Errors in Product Costing," *The Accounting Review* 69, 1994, pp. 567-591.

DEMSKI, J. *Managerial Uses of Accounting Information.* Boston, MA: Kluwer Academic Publishers, 1994.

FELLINGHAM, J., P. NEWMAN, AND Y. SUH. "Contracts without Memory in Multiperiod Agency Models," *Journal of Economic Theory* 37, 1985, pp. 340-355.

FELTHAM, G. *Information Evaluation.* Sarasota, FL: AAA, 1972.

GIGLER, F. AND T. HEMMER. "On the Frequency, Quality, and Informational Role of Mandatory Financial Reports," *Journal of Accounting Research* 36, 1998, pp. 117-147.

GIGLER, F. AND T. HEMMER. "Information Costs and Benefits of Creating Separately Identifiable Operating Segments," *Journal of Accounting and Economics* 33, 2002, pp. 69-90.

GROSSMAN, S. AND O. HART. "An Analysis of the Principal-Agent Problem," *Econometrica* 51, 1983, pp. 7-45.

GRUNFELD, Y. AND Z. GRILICHES. "Is Aggregation Necessarily Bad," *Review of Economics and Statistics* 42, 1960, pp. 1-13.

HOLMSTROM, B., "Moral Hazard and Observability," *The Bell Journal of Economics* 10, 1979, pp. 74-91.

IJIRI, Y. “Fundamental Queries in the Aggregation Theory,” *Journal of American Statistical Association*, December 1971, pp. 766-782.

IJIRI, Y. *Theory of Accounting Measurement.* Sarasota, FL: AAA, 1975.

LIANG, P. "Accounting Recognition, Moral Hazard, and Communication," *Contemporary Accounting Research* 17, 2000, pp. 457-490.

LIM, S. AND S. SUNDER. "Efficiency of Asset Valuation Rules under Price Movement and Measurement Errors," *The Accounting Review* 66, 1991, pp. 669-693.

SAPPINGTON, D., “Commitment to Regulatory Bureaucracy,” *Information Economics and Policy*, 1986, pp. 243-258.

SUNDER, S. *Theory of Accounting and Control.* Cincinnati, OH: International Thomson Publishing, 1997.