UNIVERSITÀ DEGLI STUDI DI PADOVA

Dipartimento di Scienze Economiche ed Aziendali "Marco Fanno"

THE VALUE OF INFORMATION DISCLOSURE UNDER LOCAL LEARNING. THE CASE OF FIXED TYPES

OTTORINO CHILLEMI
University of Padova

BENEDETTO GUI
University of Padova

LORENZO ROCCO
University of Padova

May 2013

"MARCO FANNO" WORKING PAPER N.161

The value of information disclosure under local learning. The case of fixed types.

Ottorino Chillemi *, Benedetto Gui **, and Lorenzo Rocco ***

Abstract

A large population plays a two-period sequential common agency game. Agents are long lived, while principals are short lived. Preferences and technology are additively separable in time and time independent. At the onset, agents are matched in pairs under private information of individual types. At the end of the first period, in each pair the principal can disclose members' reports, in which case members remain together in the second period, or conceal information, in which case members are randomly rematched and in the second period their type remains private information. We show that an equilibrium exists in which information disclosure is efficiency enhancing. Remarkably, information disclosure would have zero value if reassembling agent pairs was not an option, as in the standard one agency literature.

* ottorino.chillemi@unipd.it

Department of Economics, University of Padua, Italy.

**benedetto.gui@unipd.it

Department of Economics, University of Padua, Italy.

*** lorenzo.rocco@unipd.it

Department of Economics, University of Padua. Italy.

1. Introduction

Extracting private information is costly when the informed party can gain from disguise. An interesting question is whether information costs can be smaller in a dynamic than a static context. As is well-known this possibility has severe limitations, because of the informed party's rational reaction to the prospective reduction in his/her information rent (this problem was first studied in the ratchet effect literature: see Weitzman 1980; Holmstrom 1982; Freixas et al. 1985).

The repeated game framework is particularly clear-cut. In the standard agency game, a principal and an agent commit to a long-term contract under private information on agent type. When the type is constant over time and the objective functions of both the principal and the agent are additively separable across periods, the principal's equilibrium strategy does not use information elicited in previous periods to customize the contract: the optimal contract is the repetition of the optimal static contract (Baron and Besanko 1984; Laffont and Tirole 1988).

Using previously elicited information can also be suboptimal in other contexts. In the sequential common-agency game, two short-lived principals contract sequentially with the same long-lived agent. The upstream principal can disclose information strategically to the downstream principal. Under some assumptions—the most important being that agents' utility is separable in period-one and period-two contractual decisions—the upstream principal does not disclose information at equilibrium (see Calzolari and Pavan 2006, where an extensive review of the literature is also provided).[1]

In this paper, we study a two-period sequential common agency game involving a large population of heterogeneous fixed-type agents and, in each period, a large number of short-run agency relationships each involving two agents of privately observed types. Similarly to standard sequential common agency games, each first-period principal can disclose or conceal information on agents' reports to the second-period principal. The innovation is that we assume that the agents in a pair remain together in the second period if and only if their reports are revealed in which case these

[1] Disclosing information is suboptimal if (a) the upstream principal is not personally interested in the decisions taken by the downstream principal; (b) the agent's type is such that the sign of the single crossing condition is the same for upstream and downstream decisions; and (c) preferences in the downstream relationship are additive separable in the two contractual decisions (Calzolari and Pavan 2006, p. 169).

become common knowledge in the agency relationship; instead the pair is randomly rematched if information is not revealed. In this way, learning is local, that is the information acquired in the first period does not leak outside the group in which it has been extracted.

This framework can be applied to a variety of situations. An example is consumption of local public goods by pairs of agents with benevolent principals who maximize the joint surplus of their agent pair. The example we adopt in the remaining of the paper is subcontractors performing a joint productive activity. If perfect competition prevails among principals and some suitable assumptions hold, they end up maximizing agent surplus. Thus, since principals are surplus maximizers, we shall focus on social efficiency.

To begin with, we assume that the surplus of the productive activity is proportional to the number of high-type agents in the pair. We show that an equilibrium always exists in which information disclosure is efficiency enhancing. For some parameter values a second equilibrium exists in which information disclosure has no impact on efficiency. The difference between these equilibria is that information disclosure is selective in the former but not in the latter. Next, we show that these results are robust to some departures from the assumption of proportionality, but can fail to hold if joint surplus is less than proportional to the number of high-type agents.

In the literature on sequential common agency with many agents, Taylor (2004) considers a monopolist who can sell the list of customers to a monopolist in another market. When customers are not naive and anticipate the principal's policy, it is never profitable to the former monopolist to disclose the characteristics of his/her customers. Similar results are obtained in Dodds (2002), Ben-Shoham (2005), and Acquisti and Varian (2005) (see also Section 5 of Fudenberg and Villas-Boas 2005 for a synthesis). Our framework differs from this literature in two respects. First, an agent's payoff depends upon the type of his mates; second, type profiles of agent pairs can be distributed differently in the two periods, thanks to mobility.

The recent literature on dynamic mechanism design considers an environment in which the per-period state distribution depends on past decisions and states, as in our model. Although the treatment is more general than ours, a single agency relationship is considered (see Pavan et al., 2012; Athey and Segal, 2012).

The literature on the ratchet effect has given attention to problems that are reminiscent of ours. Kanemoto and Mac Leod (1992) study a situation in which piece-rate contracts are efficient either if parties can commit to staying together for two periods or if a competitive market for senior workers is present, since then incumbent firms cannot exploit to workers' damage the information inferred from first-period production (see Charnes et al. 2010 for experimental evidence consistent with these theoretical findings). Our framework is different because the information extracted in the first period has an efficiency-enhancing role to play.

The organization of the paper is as follows. Section 2 contains the model. Section 3 presents the main results. Section 4 discusses extensions, and Section 5 concludes. Most proofs are gathered in the Appendix.

2. The model

The game lasts two periods and is played by a countably infinite population of agents. A share $p \in]0,1[$ of them have high ability (high type), while the rest have low ability (low type). At the onset, agents are unaware of their type. Nature moves first and randomly matches pairs of agents with principals, who are drawn from a countably infinite homogeneous population. In the first period, in each agency relationship, a principal can engage agents in a sophisticated productive activity – that generates a strictly positive surplus except when both agents are low type –, or else engage them in a standard productive activity that yields zero surplus whatever agents' types. The principal proposes a revelation mechanism, to be played after agents will learn their type. The mechanism determines, as a function of each agent's confidential report of his own type: (a) whether agents' reports will be disclosed—in which case agents stay together in the second period and in the new agency relationship types are common knowledge —or instead, information will not be disclosed, in which case agents are randomly and independently rematched and in the new agency relationships their type is private information, (b) whether the high-quality or the standard activity will be carried out, and (c) transfers to the agents. In the second period, in each agency relationship, only decisions (b) and (c) have to be made. To this end, in case agents' types have to be reported again, a mechanism is implemented.

For the sake of simplicity we assume:

(i) all players' stage surplus must be nonnegative (ex-post participation constraint in each period);

(ii) each principal is a surplus maximizer;

(iii) high-type agents alone have an incentive to lie;

(iv) the surplus of the sophisticated productive activity is proportional to the number of high-type agents in the pair.

Assumption i) is sufficient to conclude that, if the standard activity is carried out at any stage, each player must obtain zero surplus in that stage. If the sophisticated activity is carried out, given assumptions (i)-(iii) joint surplus is allocated to high-type agents alone. Assumption (iv) frees us from the question of which matchings are preferable from the productive viewpoint (the surplus generated by two high types and two low types is the same regardless of whether they form two heterogeneous or two homogeneous pairs). Under these assumptions, a first-period principal has no incentive to misreport to any second-period principal.[2] Moreover, agents' stage payoffs are very simple: only high-type agents can obtain a positive payoff; this occurs when the sophisticated activity is carried out; such payoff can take only two values, associated respectively to correct and false principal's beliefs as to their types. For convenience we normalize the latter to 1 and let h denote the former. According to assumption (iii) it is $0 < h < 1$.

The following example illustrates the structure of stage payoffs. The standard activity yields an output worth .4 at a cost .2 to each subcontractor, whatever their types. The only way to satisfy participation constraints is that each agent be paid .2, so each player's surplus is 0. The sophisticated activity yields an output worth 1, at a cost .6 to a low-type subcontractor; instead, for a high-type subcontractor the cost is .3 if the other is a low and .4 if the other too is a high.[3] Therefore, when both are believed to be low type the principal cannot but choose the standard activity. In the other cases the principal can also choose the sophisticated activity, in which case the optimal payments to subcontractors are the following: if one is believed to be low type and the other high type, the former is paid .6, while the latter is paid .4; if instead both are believed high, each is paid .5. Thus a true low obtains a surplus equal to 0 anyway, while a true high obtains a surplus equal to .1 whenever the sophisticated activity is carried out. A low who is believed to be high obtains a negative payoff if the sophisticated activity is chosen, so we must not worry that he may prefer to lie. Instead, a high type can benefit from being believed to be low.

[2] When he discloses information, efficiency always obtains in the second period, so no improvement can be obtained through lying.

[3] In order for pair surplus to be proportional to the number of high types in it, some form of congestion is needed, e.g. in the use of common tools.

This occurs when the other is believed to be high and the sophisticated activity is carried out: his surplus turns out to be .2, instead of .1.[4] However, the opposite occurs when the other is believed to be low: the high type believed to be low obtains 0 for sure, since the principal certainly chooses the standard activity, while were he correctly believed high he would obtain a surplus equal to .1 , provided the principal were to choose the sophisticated activity. (The complete payoff matrix can be found in footnote.[5]) After normalization, in this example it is $h = 1/2$.

Under assumptions (i)-(iv), in each agency relationship reported type profiles are simply described by the number of reported high-type agents $i = 0, 1, 2$. A first-period principal's strategy can be described by two mappings from the number of reported high-type agents to the real numbers:

a) probabilities $\left\{r_i^1\right\}_{i=0,1,2}$ that the sophisticated activity is carried out in the first period,

b) probabilities $\{\sigma_i\}_{i=0,1,2}$ that individual reports are disclosed.

Indeed, part c) of the strategy, that is individual transfers, is trivially determined as explained above and subsumed in the payoffs. Therefore, let $\Sigma_k = (\left\{r_i^1\right\}_{i=0,1,2}, \{\sigma_i\}_{i=0,1,2})$ denote the strategy of the principal in the k-th the first-period agency relationship.

At the beginning of the second period, all agents whose principals have not disclosed agent reports are randomly assigned to brand-new agency relationships in which agents' types are private information. (A notable exception, which we will examine below, is when all other first-period principals choose $\sigma_i = 1$, all i; in this case, even when information is not disclosed players learn something about agents' types.)

[4] Notice that the mate of a high-type agent who gets a benefit from being believed to be low obtains a nonnegative payoff (namely 0) - otherwise an ex post participation constraint would be violated, and therefore such benefit could not come true, so no temptation to lie would exist.

[5] The matrix below displays an agent's stage payoff - payment received minus cost - in case the sophisticated activity is carried out, as a function of both his true and believed types, and the opponent's true type (as usual we only consider the possibility that one player at a time can lie). The couple (n, m), with $n, m \in \{low,\ high\}$, indicates that the agent is of type n and is believed to be of type m.

	$(high, high)$		(low, low)	
$(high, high)$	$.5 - .4 = .1$	$; 5 - .4 = .1$	$.4 - .3 = .1$	$; .6 - .6 = 0$
$(high, low)$	$.6 - .4 = .2$	$; .4 - .4 = 0$	$.5 - 3 = .2$	$; .5 - .6 = -.1$
(low, low)	$.6 - .6 = 0$	$; .4 - .3 = .1$	$.5 - .6 = -.1$	$; .5 - .6 = -.1$
$(low, high)$	$.5 - .6 = -.1$	$; .5 - .3 = .2$	$.4 - .6 = -.2$	$; .6 - .6 = 0$

We do not need to identify specific agency relationships within both the brand-new or the non-reassorted categories, since principals have a unique best strategy for each category. In a brand-new agency relationship, the beliefs regarding an agent's type held by other players are formulated starting from the common prior, account being taken of first-period principals' strategy profile $[\Sigma_k]_{k=1,..,..}$— where k indexes agencies. A second-period principal's strategy in a brand-new agency is denoted by Σ^f. For each first-period strategy profile,$[\Sigma_k]_{k=1,..,..}$, Σ^f consists of a mapping from the number of reported high-type agents to the real numbers, namely, probabilities $\left\{r_i^f\right\}_{i=0,1,2}$ that the sophisticated activity is carried out. In a non-reassorted agency relationship, instead, first period reports are common knowledge and a second round of reporting is not required. A principal's strategy in a non-reassorted agency relationship is denoted by $\Sigma^s = \{r_i^s\}_{i=0,1,2}$.

Finally, we assume a unit discount factor, and limit attention to symmetric equilibria.

3. Analysis

To make the problem interesting, we make the following technical assumption

(v) Inefficiency prevails in the one-period static version of the game.

Assumption (v) holds iff $p > h$. To see this, let r_i^0 , $i = 0, 1, 2$ denote the probabilities that the sophisticated activity is carried out in such game. Then, satisfaction of the incentive compatibility constraint for a high-type agent implies $phr_2^0 + (1-p)hr_1^0 - pr_1^0 - (1-p)r_0^0 \geq 0$. Efficiency requires $r_0^0 = 0, r_1^0 = 1$, and $r_2^0 = 1$, in which case the l.h.s. becomes $h - p$, which is negative under (v). QED

Now we characterize the solution for all values of p and h in the set $\{(p, h) \mid h < p < 1, h \in (0, 1)\}$. Henceforth, starred symbols will denote optimal values and symbols with superscript e will denote equilibrium values.

To begin with, we notice three simple facts.

FACT A. In the two-period version of the game, $r_0^k = 0$; $r_2^k = 1$, for $k = 1, f, s$ are always optimal: carrying out the standard activity when both agents generate zero surplus, and carrying out the sophisticated activity when both agents generate positive surplus, is efficient and encourages truthtelling. Thus, in the following, we no more discuss such probabilities.

FACT B. In the two-period version of the game, players' belief that an agent is high type in a reassorted agency relationship is $\Pi\left(\bar{\Sigma}\right)=\frac{2p^2(1-\bar{\sigma}_2)+2p(1-p)(1-\bar{\sigma}_1)}{2p^2(1-\bar{\sigma}_2)+4p(1-p)(1-\bar{\sigma}_1)+2(1-p)^2(1-\bar{\sigma}_0)}$, where $\bar{\Sigma}$ is the strategy all first-period principals play, with the possible irrelevant exception of a deviant one, provided $(\bar{\sigma}_1,\bar{\sigma}_2,\bar{\sigma}_3)\neq(1,1,1)$. In the following we will write Π instead of $\Pi\left(\bar{\Sigma}\right)$, when confusion does not arise. When instead $(\bar{\sigma}_1,\bar{\sigma}_2,\bar{\sigma}_3)=(1,1,1)$, a first-period principal cannot have his agents randomly rematched and his decision problem needs a separate analysis.

FACT C. In the one-period static version of the game it is $\left(r_0^0\right)^e=0$ and $\left(r_2^0\right)^e=1$: setting $r_0^0=0$ and $r_2^0=1$ is efficient and encourages truthtelling. Under (v) we are sure that $\left(r_1^0\right)^e=\frac{ph}{ph+p-h}$ (this is the value of r_1^0 that satisfies the incentive constraint as an equality, and when $p\geq h$ it is feasible as $0<\frac{ph}{ph+p-h}<1$).

Next, let us study the problem of a second-period principal in a reassorted agency relationship. He supposes truthtelling in the first-period, knows $\Sigma_k=\bar{\Sigma}$, $k=1,..,..$ and therefore, estimates the probability that an agent is high type as Π. Let $U^f=\Pi h+(1-\Pi)\,hr_1^f$ be the expected surplus of a high type when all agents report truthfully. Note that $U^f>0$, so we do not need to consider explicitly the participation constraint. Given the above, the principal's problem simplifies to

$$\text{maximize } 2\Pi U^f \text{ s. t.} \tag{1}$$

$$0\leq r_1^f\leq 1\quad, \tag{2}$$

$$U^f\geq \Pi r_1^f\quad. \tag{3}$$

Constraint (2) is a feasibility condition, and inequality (3) imposes that truthtelling be incentive compatible.

It is immediately seen that $r_1^{f*}=1$ if $\Pi\leq h$, and $r_1^{f*}=\frac{\Pi h}{\Pi-(1-\Pi)h}$ if $\Pi>h$. Let define $U^{f*}=\Pi h+(1-\Pi)\,hr_1^{f*}$.

Then, let us consider the problem of a second-period principal in a non-reassorted agency relationship. He faces a trivial problem: he supposes truthtelling in the first period, knows the reported type profile, and therefore, can choose the first best strategy: $r_0^s=0, r_1^s=1, r_2^s=1$.

Finally, we study the problem of a first-period principal. Let Φ denote the expected surplus of a high-type agent over the two periods. Moreover, let $\Sigma_2^*\left([\Sigma_k]_{k=1,..,.}\right)$ denote the optimal strategies of second-period principals, given the set of first-period principals' strategies (in practice, Σ_2^* consists of the two strategies Σ^{f*} and Σ^{s*}). The first-period principal of the agency under consideration solves the following problem for given $\Sigma_k = \bar{\Sigma}, k = 1, .., .$ and $\Sigma_2^* \left([\Sigma_k]_{k=1,...,} \right)$:[6]

$$\text{maximize } 2p\Phi \text{ s. t.} \tag{4}$$

$$0 \leq r_1^1 \leq 1; \quad 0 \leq \sigma_i \leq 1, i = 0, 1, 2 \tag{5}$$

$$\Phi \geq \Phi^d \tag{6}$$

where $\Phi = p\left[h + U^{f*} + \sigma_2 \left(h - U^{f*}\right)\right] + (1-p)[hr_1^1 + U^{f*} + \sigma_1 \left(h - U^{f*}\right)]$ is the expected surplus of a high-type agent; $\Phi^d = p\left[r_1^1 + U^{f*} + \sigma_1 \left(1 - U^{f*}\right)\right] + (1 - p)(1 - \sigma_0) U^{f*}$ is the surplus of a high-type agent in case he lies (and his mate tells the truth). At a symmetric equilibrium it is $\sigma_i = \bar{\sigma}_i$, $i = 0, 1, 2$.

4. Results

For a first-period principal, $\sigma_0 = 1$ is always optimal and $\sigma_2 = 1$ is always weakly optimal. Indeed, $\sigma_0 = 1$ encourages truthtelling at no cost, and $\sigma_2 = 1$ ensures a payoff at least equal to the alternative payoff and has a nonnegative impact on truthtelling constraint. Since we limit attention to symmetric equilibria, there are only two possible types of equilibria to consider.

a) Bad equilibrium. All agent pairs stay together in the second period: $(\sigma_i)^e = 1$, $i = 0, 1, 2$.

We will call the strategy of a first-period principal in such an equilibrium "autarchic strategy". Notice that, when all first-period principals play $\sigma_0 = 1$ and $\sigma_1 = 1$, second period principals are certain that their pair is made of two high-type agents, whatever the value of σ_2. This implies that efficiency obtains anyway but also introduces a payoff-irrelevant multiplicity at equilibrium. In the following we will disregard this multiplicity and assume for simplicity that when $\sigma_0 = \sigma_1 = 1$ it is also $\sigma_2 = 1$. By Fact A, it is $r_0^{1*} = 0$, $r_2^{1*} = 1$. Then, the incentive compatibility constraint becomes

$$2ph + (1-p)h\left(r_1^1 + 1\right) \geq p\left(r_1^1 + 1\right) \tag{7}$$

and hence satisfaction of the incentive constraint as an equality implies

$$\left(r_1^1\right)^e + 1 = \frac{2ph}{ph + p - h} \tag{8}$$

[6] The strategy chosen by the first-period principal under consideration has a negligible impact on second-period problems. Therefore, mentioning it among the determinants of second-period choices, or not, is immaterial.

Equation (8) tells us that a high-type's payoff is equal to twice his payoff in the one-period static version of the game – recall that $\left(r_1^0\right)^e = \frac{ph}{ph+p-h}$. Therefore, information disclosure brings no additional value. As feasibility requires $r_1^1 \geq 0$, this type of equilibrium can only occur if $\left(r_1^0\right)^e > 1/2$. This condition implies that the bad equilibrium only exists when $h < p \leq \frac{h}{1-h}$. The proof that this is actually an equilibrium is found in the Appendix.

As a corollary, the problem of a first-period principal who considers deviation when all other principals play the autarchic strategy is the same as when he supervises an isolated agency relationship (i.e. only one agent pair plays the game, so no rematching is possible). Therefore, also in the latter case, information disclosure has zero value and a high-type agent's payoff is equal to twice the one-period static payoff.

b) Good equilibrium. Some agent pairs are rematched: $(\sigma_0)^e = 1$, $(\sigma_1)^e < 1$, $(\sigma_2)^e = 1$.

In the Appendix, we prove that the good equilibrium always exists and that the equilibrium payoff of a high-type agent exceeds twice the payoff in the one-period static version of the game.

A numerical example follows. Let $p = \frac{1}{3}$ and $h = 1/4$. Since $p = \frac{h}{1-h}$ both types of equilibria exist. The good equilibrium has $r_0^1 = 0, r_1^1 = 1, r_2^1 = 1, \sigma_0 = 1, \sigma_1 = \frac{1}{4}, \sigma_2 = 1, r_0^s = 0, r_1^s = 1, r_2^s = 1, r_0^f = 0, r_1^f = \frac{1}{3}, r_2^f = 1, \Pi = \frac{1}{2}$. The expected surplus of an agent of high type is $\Phi = \frac{1}{3}(\frac{1}{4})(1+1) + \frac{2}{3}(\frac{1}{4})(1+\frac{1}{4}) + \frac{2}{3}\left(\frac{3}{4}\right)\left(\frac{1}{2}\frac{1}{3}\right) = \frac{11}{24}$. The bad equilibrium has $r_0^1 = 0$, $r_1^1 = 0$, $r_2^1 = 1$, $\sigma_0 = \sigma_1 = \sigma_2 = 1$, $r_0^s = 0$, $r_1^s = 1$, $r_2^s = 1$. The expected surplus of an agent of high type is $\Phi = \frac{1}{3}(\frac{1}{4})(1+1) + \frac{2}{3}(\frac{1}{4})(0+1) = \frac{1}{3}$. Notice that in twice the repetition of the static version of the game expected surplus is the same as in the bad equilibrium. Indeed, $r_1^0 = \frac{ph}{p-(1-p)h} = \frac{1}{2}$ and hence $\Phi = 2\frac{1}{3}(\frac{1}{4}) + 2\frac{2}{3}(\frac{1}{4})(\frac{1}{2}) = \frac{1}{3}$.

We summarize our results as follows.

Proposition 1 *There always exists a symmetric equilibrium in which information is partially disclosed and social surplus is bigger than twice the surplus of the one-period static version of the game (good equilibrium). Moreover, when $p \leq \frac{h}{1-h}$, there is a second symmetric equilibrium in which information is completely disclosed and has no impact on social surplus, which is equal to twice the surplus of the one-period static version of the game (bad equilibrium). No other symmetric equilibria exist.*

Corollary *In an isolated agency relationship, information disclosure is valueless: social surplus is equal to twice the surplus of the one-period static version of the game.*

Proposition 1 states that at equilibrium, social surplus is never diminished by information disclosure. Partial information disclosure always impacts surplus positively; when the percentage of high-type agents is smaller than $\frac{h}{1-h}$, another equilibrium exists in which information is completely disclosed and valueless.

Which is the driver of such results? Observe that in a dynamic setting like ours, as soon as a principal sets $\sigma_1 > 0$, a high-type agent has an additional temptation to behave opportunistically, with respect to the repetition of the static version of the game. Suppose he lies in the first period. If his mate is high type and the agent pair is not dissolved (an event of probability $p\sigma_1 > 0$ from his standpoint), then in the second period, he obtains 1 with certainty rather than with probability $pr_1^{0*} < 1$ (as it would occur in the repetition of the static version of the game). On the other hand, as soon as $\sigma_0 > 0$, a high-type agent gets an additional disincentive from lying in the first period. Assume he lies. If his mate is a low type and the pair is not dissolved (an event of probability $(1-p)\,\sigma_0 > 0$), then in the second period, he obtains 0 with certainty, rather than 0 with probability $1-p+p\left(1-r_1^{0*}\right)$ and 1 with probability pr_1^{0*} (as it would occur in the repetition of the static version of the game).

It turns out that there exist equilibria in which high-type agents have stronger incentives for truthful reporting than in the one-period static version of the game. Indeed, the option of keeping the pair together when both report low type, is the key to our result. It represents an especially effective punishing tool, which is absent in the static version of the game. Consider a good equilibrium. From the point of view of a high-type agent who reports low type in both periods, the expected number of times that over the two periods he meets a low type (which yields him zero, as the principal would order the standard activity) is $(1-p)(1+\sigma_0)+[(1-p)(1-\sigma_0)+p(1-\sigma_1)]\,(1-\Pi)$, which, since $\sigma_0 = 1$ and $\sigma_1 < 1$ (and, therefore, $\Pi < 1$) is always greater than twice the expected number of times in the one-period static version of the game, that is, $2(1-p)$. By the same token, in a bad equilibrium, a high-type agent who falsely reported low type in the first period does not have to report again in the second period, and hence, the expected number of times that he meets a low type over the two periods equals exactly $2(1-p)$. The lesson is that the potential favorable dynamic effect of disclosure only occurs if disclosure is selective. Indeed, selectivity opens the possibility of unfavorable matchings to misreporting agent in the second period. The same line of reasoning provides an intuition of why information disclosure has zero value in the case of an isolated agency relationship.

5. Extensions

The extension to agencies with more than two agents does not seem to offer new perspectives on information disclosure. More interesting is the question of what can we learn if we remove the assumption that the surplus of the sophisticated productive activity is proportional to the number of high-type agents in the pair. Suppose that assumption (iv) is substituted by the following:

(iv*) when the sophisticated activity is carried out the stage payoff of a truthful high type agent is $h > 0$ if his mate is high type, $z > 0$ if his mate is low type, and $H > 0$ if he falsely reports low type.

Now, a first-period principal faces the additional problem of optimally matching each of his agents with a second-period mate. The analysis of such model is complicated because the share of high type in the population of agents to be rematched in the second period is a non linear function of the strategy played by first-period principals. However, it is possible to show that the following results hold.

Proposition 2

Under assumptions (i), (ii), (iii), (iv), (v)*

a) The equilibrium expected payoff of a high-type agent is no lower than twice the payoff he could get in the corresponding one-shot static game

b) There is a non negligible set of parameter values in which the bad equilibrium is the only equilibrium

c) There is a nonnegligible set of parameter values in which only one equilibrium of the good type exists.

d) There is a nonnegligible set of parameter values in which both good and bad equilibria exist.

In the Appendix we provide a proof of part b) of Prop 2 – the complete proof can be obtained from the authors.

The new result is that a set of parameters exists in which information disclosure is always complete and therefore valueless at equilibrium. The intuition of this result is that when $H < z$ and the autarchic strategy is feasible, it is in the individual interest of first-period principals to set $\sigma_0 = \sigma_1 = 1$, while keeping the incentive constraint satisfied by setting r_1^1 at the suitable level. The equilibrium strategy is autarchic, and yields twice the one-shot static payoff. Given these conditions on parameter values, however, it can be shown that feasible strategies with $\sigma_1 < 1$ exist such that the payoff would be greater, but this opportunity is wiped out by the free riding behavior of first-period principals.

6. Conclusion

The difficulty of extracting information to customize contracts in a long-term relationship lies in the fact that the informed party has more opportunities to lie than in a one-shot interaction.

It is this state of things that sustains the bad equilibrium, the one in which no agent pair is rematched. Consider for instance the possible deviation of a first-period principal who, instead of always disclosing information, conceals it with positive probability when the reported number of high types is one. At this point, a high-type agent who has lied in the first period, if required to report again in the second period, is certain that his mate is of high type, so he will optimize his report and lie again. It is the cost of preventing such a chain of lies that prevents players from improving upon the repetition of the static equilibrium.

Instead, when at equilibrium a positive share of agent pairs are rematched, information disclosure has social value. The key to this result is that, thanks to rematching, some of the information locally learned by agents can be concealed by their principal. Then new ways of punishing agent deviations become available at equilibrium, that are not available in a static setting.

Appendix

Part A) We prove that bad equilibria exist iff $p < \frac{h}{1-h}$.

When $\bar{\sigma}_i = 1$, all i, a first-period principal who considers deviation solves the following:

$$\text{maximize} \quad 2p\left\{ph(r_2^1+\sigma_2)+(1-p)h\left[r_1^1+\sigma_1+(1-\sigma_1)r_1^{f*}\right]\right\} \text{ s. t.}$$

$$ph(r_2^1+\sigma_2)+(1-p)h\left[r_1^1+\sigma_1+(1-\sigma_1)r_1^{f*}\right] \geq \tag{9}$$
$$p\left(r_1^1+\sigma_1\right)+\max\left[p(1-\sigma_1)h+(1-p)(1-\sigma_0)hr_1^{f*}, p(1-\sigma_1)r_1^{f*}\right]$$

where r_2^{s*} and r_2^{f*} have been set to their optimal value 1, and r_1^{f*} is optimally chosen by second-period principals, given the strategy of first-period principals. Hence, $r_1^{f*} \in [0,1]$ is the maximum in the set of solutions of the following second-period incentive constraint (that is designed for an agent who told the truth in the first period):

$$p(1-\sigma_2)h+(1-p)(1-\sigma_1)hr_1^f \geq p(1-\sigma_2)r_1^f. \tag{10}$$

First, notice that setting $r_2^1=1$ is optimal since it encourages truthtelling and maximizes the objective. Also, setting $\sigma_0=1$ is optimal, since it weakly encourages truthtelling and does not affect the objective function. Then, setting $\sigma_2=1$ is optimal since it encourages truthtelling and maximizes the objective—notice that setting $\sigma_2=1$ makes one sure that no agent has incentive to lie in the second period if he did not lie in the first — constraint (10) becomes $(1-p)(1-\sigma_1)hr_1^f \geq 0$, so $r_1^{f*}=1$.

When $\sigma_0=1$ is announced, the incentive compatibility constraint (9) tells us that a high-type agent who is considering of falsely reporting low type in the first period knows that, in the event that he is asked to report again in the second period, his partner is high type. Accordingly, he will plan to lie in the second period conditionally on having lied in the first. Substituting the optimal values determined above, the first-period principal's problem boils down to the following:

$$\text{maximize}_{\,r_1^1}\ 2p\left\{2ph+(1-p)h\left[r_1^1+1\right]\right\}\ \ s.t. \tag{11}$$

$$2ph+(1-p)h\left[r_1^1+1\right] \geq p\left[r_1^1+1\right] \tag{12}$$

The value of the problem does not depend on σ_1, and hence $\sigma_1=1$ is optimal. Lastly (12) is satisfied by $r_1^1 \geq 0$ iff $p \leq \frac{h}{1-h}$. QED

Part B) We prove that the good equilibrium always exists and the equilibrium payoff of a high-type agent exceeds twice the payoff in the one-period static version of the game.

First, observe that $\bar{\sigma}_0=1$, $\bar{\sigma}_1<1, \bar{\sigma}_2=1$ implies $\Pi=1/2$ independently of the value of $\bar{\sigma}_1$ in the $[0,1[$ interval. This allows us to focus on the maximization problem of individual principals, since they do not affect each other through their choices. There are two cases to consider.

Case 1. $h \geq 1/2$.

Then, $\Pi = 1/2 \leq h$ and hence it is immediately seen that $r_1^{f*} = 1$ and $U^{f*} = h$. The incentive constraint writes

$$ph(1+1) + (1-p)h\left(r_1^1 + 1\right) \geq p(r_1^1 + \sigma_1) + p\left(1 - \sigma_1\right)h. \tag{13}$$

Setting $\sigma_1 = 0$ is always optimal for a single principal, since it encourages truthtelling at no cost (in reassorted agency relationships, efficiency prevails as well). If the incentive constraint (13) is not binding, then $r_1^{1*} = 1$: the dynamic equilibrium is fully efficient, so the thesis trivially holds. If the incentive constraint is binding, it writes $ph(1+1)+(1-p)h\left(r_1^1 + 1\right) = pr_1^1 + ph$, hence $r_1^{1*} = \frac{h}{p-(1-p)h} > \frac{ph}{p(1+h)-h} = r_1^{0*}$.

Since it is $r_1^{k*} = 1 > r_1^{0*}$ for $k = s, f$, the dynamic equilibrium payoff is surely bigger than twice the payoff of the one-period static version of the game.

Case 2. $h < 1/2$

Then, $\Pi = 1/2 > h$, so inefficiency prevails in reassorted agency relationships; substitution into the formula $r_1^{f*} = \frac{\Pi h}{\Pi-(1-\Pi)h}$ yields therefore $r_1^{f*} = \frac{h}{1-h}$ and therefore $U^{f*} = \frac{h}{2}(1 + r_1^{f*}) = \frac{h}{2(1-h)}$. The incentive constraint writes

$$\Phi = ph(1+1) + (1-p)h\left(r_1^1 + \sigma_1\right) + (1-p)\left(1 - \sigma_1\right)U^{f*} \geq p(r_1^1 + \sigma_1) + p\left(1 - \sigma_1\right)U^{f*} = \Phi^d. \tag{14}$$

Now note the following facts:

a) $\frac{\partial 2p\Phi}{\partial \sigma_1} = 2p(1-p)(h - U^{f*}) > 0$.

b) $\frac{\partial(\Phi-\Phi^d)}{\partial\sigma_1} = (1-p)(h - U^{f*}) - p\left(1 - U^{f*}\right) < (1-h)(h - U^{f*}) - h\left(1 - U^{f*}\right) = -\left(h^2 + (1-2h)U^{f*}\right) < 0$, where the former inequality descends from $p > h$ and the latter from $h < 1/2$.

c) $-\frac{\partial 2p\Phi}{\partial r_1^1} / \frac{\partial(\Phi-\Phi^d)}{\partial r_1^1} = -\frac{2p(1-p)h}{(1-p)h-p} > 0$, $-\frac{\partial 2p\Phi}{\partial\sigma_1} / \frac{\partial(\Phi-\Phi^d)}{\partial\sigma_1} = -\frac{2p(1-p)(h-U^{f*})}{(1-p)(h-U^{f*})-p\left(1-U^{f*}\right)} > 0$, $-\frac{\partial 2p\Phi}{\partial\sigma_1} / \frac{\partial(\Phi-\Phi^d)}{\partial\sigma_1} < -\frac{\partial 2p\Phi}{\partial r_1^1} / \frac{\partial(\Phi-\Phi^d)}{\partial r_1^1}$ for any $U^f > 0$.[7]

Notice that $-\frac{\partial(2p\Phi)}{\partial x} / \frac{\partial(\Phi-\Phi^d)}{\partial x}$ is the absolute value of the decrease in the objective function per unit increase in the truthtelling constraint obtained by lowering the variable x, where $x = r_1^1$, σ_1 – assumption (v) ensures that the denominator is negative.

[7] In fact $\frac{2p(1-p)(1-U^{f*})}{p\left(1-U^{f*}\right)-(1-p)(h-U^{f*})} = \frac{2p(1-p)}{p-(1-p)\left(\frac{h-U^{f*}}{1-U^{f*}}\right)} < \frac{2p(1-p)}{p-(1-p)h}$ as $h < 1$, $h > U^{f*}$,and $p - (1 - p)\left(\frac{h-U^{f*}}{1-U^{f*}}\right) > 0$.

Fact c) means that lowering σ_1 to satisfy the incentive constraint is more cost-effective than lowering r_1^1.

At equilibrium, two cases are possible:

i) $\sigma_0 = \sigma_2 = 1, 0 < \sigma_1 < 1$, and $r_1^1 = 1$

ii) $\sigma_0 = \sigma_2 = 1, \sigma_1 = 0$, and $0 < r_1^1 < 1$.

(When both r_1^1 and σ_1 are at 0 (14) is satisfied as a strict inequality.)

Last, we prove that in both cases at equilibrium the payoff is bigger than the repetition of the one-shot static payoff, that is the following holds

$$2ph + (1-p)(hr_1^1 + U^{f*}) - 2[ph + (1-p)hr_1^0] > 0. \tag{15}$$

Case i) (14) as an equality yields the equilibrium value of σ_1. From

$$ph(1+1)+(1-p)h(1+\sigma_1)+(1-p)(1-\sigma_1)\frac{h}{2(1-h)} = p(1+\sigma_1)+p(1-\sigma_1)\frac{h}{2(1-h)}$$

we get $\sigma_1 = \frac{p-2hp-h(-p+1)+h\frac{p}{-2h+2}-\frac{h}{-2h+2}(-p+1)}{-p+h(-p+1)+h\frac{p}{-2h+2}-\frac{h}{-2h+2}(-p+1)}$. Then, (15) becomes

$$(-2)\left(h-2p+2hp-2h^2+2h^2p\right)^{-1}(p-h+hp)^{-1}(p-h)(p-1)(h-1)hp > 0$$

Notice that for $0 < h < 1/2$ and $h < p$ the term $h - 2p + 2hp - 2h^2 + 2h^2p$ is always negative, since it is increasing in h and negative at its maximum for each given $p > h$.

Case ii) (14) as an equality yields the equilibrium value of r_1^1. From

$$ph + (1-p)hr_1^1 + p(h - U^{f*}) + (1-p)U^{f*} = pr_1^1. \tag{16}$$

we get $r_1^1 = \frac{h}{2}\frac{1+2p-4hp}{(p-(1-p)h)(1-h)}$. Then, (15) becomes

$$\frac{1}{2}h\frac{1+2p-4hp}{(p-(1-p)h)(1-h)} + \frac{1}{2}\frac{1}{1-h} - \frac{2ph}{p-(1-p)h} = \frac{1}{2}\frac{p}{p-(1-p)h} > 0$$

since $p > h$. QED

Part C) We prove Proposition 2 part b).

Suppose $0 < h < H < z = 1$ after normalization. Assumption (v) implies $ph + (1-p) < pH$, i.e. $p > \frac{1}{H-h+1}$. Observe that it is $\frac{1}{H-h+1} > 1/2$

We will show that the bad equilibrium is the unique equilibrium in a subspace of parameters. The fact that the bad equilibrium exists when the autarchic strategy is feasible is easily seen by the same reasoning we used above. Therefore existence is proved if we verify that the autarchic strategy is feasible. The incentive compatibility constraint of a high-type agent writes

$$\Phi = ph(1+\sigma_2) + p\,(1-\sigma_2)\,U^{f*} + (1-p)(r_1^1 + \sigma_1 + (1-\sigma_1)\,U^{f*}) \geq \tag{17}$$
$$\left(pH(r_1^1+\sigma_1) + p\,(1-\sigma_1)\,U^{f*} + (1-p)\,(1-\sigma_0)\,U^{f*}\right) = \Phi^d.$$

which for $\sigma_i = 1, i = 0,1,2$ reduces to

$$2ph + (1-p-pH)(1+r_1^1) \geq 0. \tag{18}$$

Since assumption (v) implies $pH > (1-p)$, the inequality is satisfied by $r_1^1 \geq 0$ iff $\frac{2ph}{pH-(1-p)} - 1 \geq 0$, i.e. $h \geq \frac{pH-(1-p)}{2p}$, which characterizes a non empty subset of the parameter space (observe that it is $\frac{1}{2} \geq \frac{pH-(1-p)}{2p} > 0$).

Lastly, let us prove that the bad equilibrium is the unique equilibrium when it exists.

First notice that the following facts hold true:

Fact 1. When $\sigma_0 = 1$ it is certainly $\frac{\partial 2p\Phi}{\partial \sigma_1} = 2p(1-p)(1-U^{f*}) > 0$ (in fact $U^{f*} = \Pi h + (1-\Pi)\,r_1^{f*} < 1$ whenever $\Pi > 0$).

Fact 2. When $\frac{\partial\left(\Phi-\Phi^d\right)}{\partial\sigma_1} = (1-p)(1-U^{f*}) - p\left(H-U^{f*}\right) < 0$ it is $-\frac{\partial(2p\Phi)}{\partial\sigma_1}/\frac{\partial\left(\Phi-\Phi^d\right)}{\partial\sigma_1} > -\frac{\partial(2p\Phi)}{\partial r_1^1}/\frac{\partial\left(\Phi-\Phi^d\right)}{\partial r_1^1} > 0$, as $-\frac{2p(1-p)(1-U^{f*})}{(1-p)(1-U^{f*})-p\left(H-U^{f*}\right)} > -\frac{2p(1-p)}{(1-p)-pH}$.[8]

Now it is clear that a good equilibrium - that would have $(\sigma_1)^e < 1$ - cannot exist: the maximand increases monotonically in σ_1, account being taken of the incentive constraint. In fact, if $\frac{\partial\left(\Phi-\Phi^d\right)}{\partial\sigma_1} \geq 0$ increasing σ_1 beyond $(\sigma_1)^e$ also loosens the constraint; if instead $\frac{\partial\left(\Phi-\Phi^d\right)}{\partial\sigma_1} < 0$ fact 2 ensures that increasing σ_1 beyond $(\sigma_1)^e$ and at the same time lowering r_1^1 so as to keep the constraint binding, raises the maximand. As we know by assumption that this move is feasible, no good equilibrium can exist. QED

References

Acquisti, Alessandro, and Hal R.Varian. 2005. "Conditioning Prices on Purchase History." Marketing Science 24(3): 367-381.

[8] In fact $\frac{2p(1-p)(1-U^{f*})}{p\left(H-U^{f*}\right)-(1-p)(1-U^{f*})} = \frac{2p(1-p)}{p\left(\frac{H-U^{f*}}{1-U^{f*}}\right)-(1-p)} > \frac{2p(1-p)}{pH-(1-p)}$ as $H < 1$ and $p\left(\frac{H-U^{f*}}{1-U^{f*}}\right) - (1-p) > 0$.

Athey, Susan, and Ilya Segal. 2012 "An Efficient Dynamic mechanism", working paper Stanford University.

Baron, David P., and David Besanko. 1984. "Regulation and Information in a Continuing Relationship." Information Econ. and Policy 1: 267-330.

Ben-Shoham, Assaf. 2005. "Information in Sequential Trade." Manuscript. Harvard Univ.

Calzolari, Giacomo, and Alessandro Pavan. 2006. "On the optimality of privacy in sequential contracting." J. Econ. Theory 130 (1): 168-204.

Charness, Gary, Peter Kuhn, and Marie Claire Villeval. 2011. "Competition and the Ratchet Effect." J. Labor Econ. 29: 513-547.

Dodds, Stefan. 2002. "Confidentiality Rules and Welfare: A Dynamic Contracting Approach with Two Principals." Manuscript. Queen's Univ.

Freixas, Xavier, Roger Guesnerie, and Jean Tirole. 1985. "Planning under Incomplete Information and the Ratchet Effect." Rev. Econ. Studies 52: 173-91.

Fudenberg, Drew, and J. Miguel Villas-Boas. 2006. "Behavior-Based Price Discrimination and Customer Recognition." in Economics and Information Systems, edited by Terrence Hendershott. Amsterdam: Elsevier.

Holmstrom, Bengt. 1982. "Design of incentive schemes and the new Soviet incentive model." European Econ. Rev. 17: 127-148.

Kanemoto,Yoshitsugu, and W. Bentley MacLeod. 1992. "The Ratchet Effect and the Market for Secondhand Workers." J. of Labor Econ. 10: 85-98.

Laffont, Jean-Jacques, and Jean Tirole. 1988. "The Dynamics of Incentive Contracts." Econometrica 56: 1153-75.

Pavan, Alessandro, Ilya Segal, and Juuso Toikka. 2012 "Dynamic Mechanism Design", working paper Stanford University.

Taylor, Curtis R. 2004. "Consumer Privacy and the Market for Customer Information." RAND J. Econ. 35: 631-650.

Weitzman, Martin L. 1980. " The 'Ratchet Principle' and Performance Incentives." Bell J. Econ. 11: 302-08.