UNIVERSITÀ DEGLI STUDI DI PADOVA

dSEA

Ottorino Chillemi, University of Padova
Stefano Galavotti, University of Padova
Benedetto Gui, University of Padova

INEFFICIENT RATIONING WITH POST-CONTRACTUAL INFORMATION

November 2017

Marco Fanno Working Papers - 214

Inefficient rationing with post-contractual information

Ottorino Chillemi* Stefano Galavotti† Benedetto Gui‡

Abstract

We study a contractual design problem between a seller and a buyer where some information correlated with the buyer's valuation is publicly observed ex-post and the allocation, but not payments, can be made contingent on it. Our analysis shows that, to maximize her profit, the seller should offer one contract in which the good is transferred to the buyer only if the ex-post signal turns out to be bad; this generates inefficient rationing: some buyers with low valuation are assigned the good more often than others with higher valuation. We show that, in contrast with previous results, the optimal contract may decrease social welfare relative to the case in which no signal is available (or it is not used). We apply our model to interpret two real-world situations: internet plans with bandwidth caps for mobile phones and promotion schemes in organizations with exogenously fixed wages.

JEL classification: D82, D86.
Keywords: rationing, ex-post information, mechanism design.

*Department of Economics and Management, University of Padova, Italy. Email: ottorino.chillemi@unipd.it.

†**Corresponding author.** Department of Economics and Management, University of Padova, via del Santo 33, 35123 Padova, Italy. Phone: +39 049 8274055. Fax: +39 049 8274221. Email: stefano.galavotti@unipd.it.

‡Department of Economics and Management, Istituto Universitario Sophia, Loppiano (Florence), Italy. Email: benedetto.gui@unipd.it.

1 Introduction

Several contractual relationships have the characteristic that, during the life of the contract, one party may get to know some information about the counterparty that was not available prior to contracting. For example, an employer can typically better appreciate the true skills of her new employees only after they start working on their job. Insurance companies usually only have a superficial knowledge of the risk profile of their customers at the time of providing them with insurance; the same is true for a lender about the creditworthness of the borrower. In both cases, events that occur later – e.g. an heart attack suffered by the insured, or a bad report on the future profitability of the borrowing firm – could be used to better shape the true characteristics (riskiness, profitability) of the other party.

It has long been recognized that, in these situations, the uninformed party would benefit from a contract whose provisions adapt to this delayed information, as this would permit to reduce the information rent enjoyed by the informed party.

In particular, in a principal-agent framework, Riordan and Sappington (1988) show that the arrival of even a little amount of information that reduces the information asymmetry about the agent's type may allow the principal to simultaneously maximize social surplus and leave no rent to the agent. Essentially, the contract must include a premium if the new information is coherent (in expectation) with the contractual choice of the agent and a penalty in the opposite case. Cremer and McLean (1988) and McAfee and Reny (1992) extended this intuition to a multiple agents' framework: in this case, the principal does not even need to observe a signal himself, but may exploit the correlation across agents' signals.[1]

These results have the disturbing feature that premia and penalties must be larger and larger as the signal becomes less and less informative. When constraints on the agent's admissible payments or payoffs are present, the principal can still exploit the information revealed ex-post to increase her profit, but full surplus extraction (and efficiency) will generally not be attainable. This is shown by Robert (1991) in an auction context, and by Kosmopolou and Williams (1998) in a general mechanism design problem. Demougin and Garvie (1991) prove that, under limited liability, both the principal's profit and efficiency increase with the informativeness of the signal.

The common feature of all the results described above is that the element of the contract that is contingent on the post-contractual information is the payments' schedule: by penalizing the agent when the information is incoherent with her reported type, payments allow the principal to induce truthtelling without the need of leaving much information rent to the agent. Given truthtelling, the allocation rule is then used to increase social surplus,

[1] More precisely, full surplus extraction is possible only if a certain condition relating the agents' payoff and the signal distributions are satisfied, or if the set of ex-post signals is rich enough. Bose and Zhao (2007) characterize the optimal contract when these conditions fail to hold.

that is eventually appropriated by the principal herself through the transfers themselves.

Despite the focus of the literature on the role of contingent payments, there are real-world examples where the ex-post information affects other, non-price elements of the contract. Consider home Internet service and mobile data plans with bandwidth caps. Some plans stipulate that, when the user exceeds a certain threshold on data transfer, the provider may throttle the transfer rate. In these contracts, the user pays a fixed price ex-ante, and no further contingent payments are required. On the other hand, the service itself (i.e. the internet connection speed) depends on an ex-post signal (whether or not the consumer has hit the threshold). One possible reason why, in this case, the firm prefers not to require additional future payments if the threshold is reached, may be that obtaining the agreed payment from the consumer could be difficult or very costly (especially when the customer has a rechargeable SIM card), whereas the firm can easily and immediately implement a lower connection speed.

To model situations like the one described above, we consider a seller-buyer context, where the buyer's valuation for the good on sale is private information, the seller observes a binary signal that is correlated to the buyer's valuation, and designs a contract where the allocation rule, but not payments, can be contingent on it. Our main finding is that the optimal contract has an allocation rule that is non-monotone in the buyer's type. In particular, the allocation rule works as follows. For buyers with low and high valuations, the signal is not used at all: the former are not assigned the good, the latter receive the good for sure. For buyers with intermediate valuations, instead, the allocation depends on the signal, but in a seemingly counterintuitive fashion: the good is transferred only if the signal turns out to be bad, while it is retained by the seller if the signal is good. As a consequence, within this interval of valuations, buyers are inefficiently rationed: those with lower valuations are assigned the good more often than those with higher valuations. This last fact generates per se a loss in efficiency. It is then not surprising that, as we show, the presence of post-contractual information can be harmful for social welfare. Hence, in contrast with previous results, we find that, even with linear utilities and absent any limited liability constraint, efficiency can decrease with post-contractual information.

Intuitively, inefficient rationing allows the seller to extract surplus from relatively low types, without leaving much information rent to high types. In fact, the mechanism assigns the good with high probability to relatively low types (because these types are very likely to have a bad signal): hence, the principal can extract a fair amount of surplus from them; at the same time, high types find it unattractive to underreport their types, because, if they did so, they would be assigned the good with very little probability (they are likely to have a good signal).

More generally, our model can be applied to contractual relations where, even though some information becomes available ex-post, the principal cannot (or does not want to) obtain a further performance (be it monetary or not) from the agent, but still can use it to shape other elements of the agreement. As an example, consider institutions with an administered salary system. In these organizations, it is often the case that wages associated with different levels are exogenously fixed and an employee can climb the wage ladder only through promotions (see Van Herpen et al. (2006) for clean evidence that promotions play the major role in determining individual wages). It is realistic to think that, in order to increase her chances of getting a promotion, an employee should be willing to comply with (unpaid) additional activities that go beyond her contractual duties. A situation like this can be interpreted as an implicit contract where the employee makes an ex-ante "payment" in terms of these additional activities; in exchange for this, the employer promises a promotion in the future that is conditional on the realization of a future signal correlated with the employee's type (e.g. her productivity). Applied to this context, our model would then predict that, for workers with intermediate levels of productivity, less productive ones should be promoted more often. Notice that, in this case, requiring future payments (i.e. activities) is not feasible for the principal (the employer) because, being implicit, the contract is not enforceable.

The rest of the paper is organized as follows. Section 2 outlines the model, that is solved and analyzed in Section 3. Section 4 presents the two applications described above showing how the model can easily be adapted to them. Section 5 explores the welfare implications of the optimal contract derived in Section 3. Section 6 concludes.

2 The Model

Consider the following contractual relationship between the seller of a good (the principal) and a buyer (the agent). The contract has two elements: a decision or allocation $\pi \in [0, 1]$ – whether or not the good is transferred from the seller to the buyer or, in general, the probability that this occurs – and a transfer or payment t (potentially negative) from the buyer to the seller. If the decision π and the transfer t are implemented, the buyer's payoff is $u_B = \pi \times v - t$, while the seller's payoff is simply $u_S = t$. The parameter v represents the valuation that the buyer attaches to the good and is private information. We will also call v the buyer's "type". The seller knows that v is the realization of a continuous random variable defined on the interval $V = [\underline{v}, \overline{v}]$, where $\overline{v} > \underline{v} \geq 0$, with cumulative distribution function G, and strictly positive density g.[2]

After the buyer selects a contract (if any), the seller receives a publicly observable and non-manipulable signal s, that we assume to be binary ($s \in \{0, 1\}$), and that is correlated with the buyer's type. Let $p_0(v)$ be the probability of observing signal $s = 0$

[2]Notice that the assumptions that $v \geq 0$ and that the sellers bears no cost for providing the good jointly imply that it would be efficient to always transfer the good to the buyer.

when the buyer's type is v. Clearly, for all v, $p_1(v) = 1 - p_0(v)$. We assume that $p_0(v)$ is a differentiable and strictly decreasing function of v.[3]

To better screen the buyer, the seller can use the information provided by the signal. Our crucial assumption is that the decision π can be made contingent on the realization of the signal, while the transfer t cannot.

Thanks to the revelation principle, the contract design problem can be framed in terms of the choice of a direct revelation mechanism; under our assumptions, a contract can then be summarized by a set of three functions $C(s, \hat{v}) = \{\pi_s(\hat{v}); t(\hat{v})\}_{s=0,1}$, with the following interpretation: (1) the seller commits to the mechanism $C(s, \hat{v})$; (2) the buyer accepts or reject the mechanism; (3) in case of acceptance, the buyer reports her type $\hat{v}$; (4) depending on the report, the transfer $t(\hat{v})$ is determined and executed; (5) the signal s is publicly observed; (6) depending on the realization of the signal ($s = 0$ or $s = 1$) and on the report, the allocation $\pi_s(\hat{v})$ is implemented.

For $s = 0, 1$, let $\eta_s(v) = \frac{dp_s(v)}{dv}\frac{v}{p_s(v)}$ be the type-elasticity of signal s and let $\psi_s(v) = v - \frac{1-G(v)}{g(v)}(1 + \eta_s(v))$ be the virtual valuation associated with signal s.[4] We make the following regularity assumptions:

(A1) for $s = 0, 1$, $\exists\ v_s \in (\underline{v}, \overline{v})$ such that $\psi_s(v) < 0$ for $v < v_s$, $\psi_s(v_s) = 0$, $\psi_s(v_s) > 0$ for $v > v_s$;

(A2) $p_0(v)v < p_0(v_0)v_0$ for $v < v_0$; $p_0(v)v \geq p_0(v_0)v_0$ for $v_0 \leq v \leq v_1$.

Assumption (A1) guarantees that the solution to our problem is non-trivial and well-behaved. Notice that (A1) implies that $v_0 < v_1$. Assumption (A2) is necessary to satisfy incentive compatibility.

3 Analysis

To characterize the optimal menu of contracts, thanks to the revelation principle, we can restrict our attention to truth-telling equilibria of the direct revelation mechanism $C(s, \hat{v})$.

[3]This is exactly equivalent to saying that the likelihood ratio $p_0(v)/p_1(v)$ is strictly decreasing.

[4]The virtual valuation is a concept taken from auction theory and it can be interpreted as a marginal revenue. In particular, the virtual valuation associated with signal s can be given the following interpretation: suppose the seller makes a take-it-or-leave-it offer at a price t, with the agreement that the good will be delivered only if the signal turns out to be equal to s. Under this offer, there will be a type v who is just indifferent between accepting and rejecting the offer (while all higher types will strictly prefer to accept). To increase the probability of selling the good, the seller has to decrease the indifferent type below v (through a proper reduction of t). The virtual valuation of type v is the additional revenue accruing to the seller associated to a marginal decrease of the indifferent type v.

The expected payoff of the seller if the buyer always reports truthfully is:

$$U_S = \int_V t(v)g(v)dv. \tag{1}$$

The expected payoff of the buyer, type v, if she reports $\hat{v}$ is:

$$U_B(\hat{v}; v) = E[\pi_s(\hat{v}); v]v - t(\hat{v}),$$

where $E[\pi_s(\hat{v}); v] = p_0(v)\pi_0(\hat{v}) + p_1(v)\pi_1(\hat{v})$ is the expected allocation.[5]

The optimal contract is the one that maximizes the seller's expected payoff, conditional on the buyer accepting the contract and reporting truthfully, i.e. it is the solution to the following program P:

$$\max_{C(s,v)} U_S,$$

subject to:

$$v \in \arg\max_{\hat{v}} U_B(\hat{v}; v), \quad \text{for all } v \in V, \tag{2}$$

and to:

$$U_B(v; v) \geq 0, \quad \text{for all } v \in V. \tag{3}$$

Constraint (2) is the incentive compatibility constraint, which requires that, for every buyer's type, truth-telling is a best reporting strategy. Constraint (3) is the individual rationality or participation constraint, which requires that signing the contract is convenient for all types.

Notice that, by the envelope theorem, (2) implies that

$$\left.\frac{dU_B(\hat{v}; v)}{dv}\right|_{\hat{v}=v} = \left.\frac{\partial U_B(\hat{v}; v)}{dv}\right|_{\hat{v}=v} \tag{4}$$

where

$$\left.\frac{\partial U_B(\hat{v}; v)}{dv}\right|_{\hat{v}=v} = \sum_{s=0}^{1} p_s(v)\pi_s(v)(1 + \eta_s(v)).$$

To solve for the optimal mechanism, we relax program P, by neglecting (3) and by replacing constraint (2) with the necessary condition (4). We obtain a relaxed program P', with maximand (1) and constraint (4). We will then check ex-post that the solution to P' does indeed solve P as well.

[5] Notice that the optimal mechanism is one for which the revenue equivalence theorem holds. To see this, note that payments enter linearly in the payoffs of both the principal and the agent (see, e.g., Krishna 2010, Chapter 3).

Now, the expected payoff of the seller in any incentive compatible mechanism writes

$$\int_V \left[\sum_{s=0}^{1} p_s(v)\pi_s(v)v - U_B(v;v) \right] g(v)dv.$$

Using (4) it is possible to reduce the above maximand to[6]

$$\int_V \sum_{s=0}^{1} \left[v - \frac{1-G(v)}{g(v)}(1+\eta_s(v)) \right] p_s(v)\pi_s(v)g(v)dv. \tag{5}$$

Now we are ready to state the main result.

PROPOSITION 1. *The optimal mechanism is*

$$\pi_s^*(v) = \begin{cases} 0 & \text{if} \quad \underline{v} \le v < v_s \\ 1 & \text{if} \quad v_s \le v \le \overline{v} \end{cases}, \qquad s = 0, 1; \tag{6}$$

$$t^*(v) = \begin{cases} 0 & \text{if} \quad \underline{v} \le v < v_0 \\ p_0(v_0)\, v_0 & \text{if} \quad v_0 \le v < v_1 \\ p_0(v_0)\, v_0 + p_1(v_1)v_1 & \text{if} \quad v_1 \le v \le \overline{v} \end{cases}.$$

Proof. Under Assumption (A1), the expression within square brackets in (5) is strictly negative for $v < v_s$, strictly positive for $v > v_s$. Hence, (5) is maximized pointwise by the allocation stated in (6). To pin down $t(v)$, notice that, for all v:

$$t(v) = E[\pi_s(v); v]v - U_B(v;v),$$

where

$$U_B(v;v) = \int_{\underline{v}}^{v} \frac{\partial U_B(u;u)}{\partial u} du + U_B(\underline{v};\underline{v}) = \int_{\underline{v}}^{v} \sum_{s=0}^{1} p_s(u)\pi_s(u)(1+\eta_s(u))du + U_B(\underline{v};\underline{v}).$$

It is immediate to verify that, given $\pi_s^*(v)$, $t^*(v)$ solves the above equations for all v, with $U_B(\underline{v};\underline{v}) = 0$. In fact, for $v < v_0$, we have:

$$t^*(v) = E[\pi_s^*(v); v]v - \int_{\underline{v}}^{v} \sum_{s=0}^{1} p_s(u)\pi_s^*(u)(1+\eta_s(u))du = 0;$$

for $v_0 \le v < v_1$, we have:

$$\begin{aligned} t^*(v) &= E[\pi_s^*(v); v]v - \int_{\underline{v}}^{v} \sum_{s=0}^{1} p_s(u)\pi_s^*(u)(1+\eta_s(u))du \\ &= p_0(v)v - \int_{v_0}^{v} p_0(u)(1+\eta_0(u))du = p_0(v_0)v_0; \end{aligned}$$

[6] The result follows from integration by parts.

for $v_1 \leq v \leq \overline{v}$, we have:

$$\begin{aligned} t^*(v) &= E[\pi_s^*(v); v]v - \int_{\underline{v}}^{v} \sum_{s=0}^{1} p_s(u)\pi_s^*(u)(1+\eta_s(u))du \\ &= v - \int_{v_0}^{v} p_0(u)(1+\eta_0(u))du - \int_{v_1}^{v} p_1(u)(1+\eta_1(u))du = p_0(v_0)v_0 + p_1(v_1)v_1. \end{aligned}$$

This shows that $(\pi_s^*(v), t^*(v))$ solve program P'. To see that they also solve program P, we first have to check that (2) is satisfied. Now, under $(\pi_s^*(v), t^*(v))$, we have that:

$$\begin{aligned} U_B(\underline{v} \leq \hat{v} < v_0; v) &= 0, \\ U_B(v_0 \leq \hat{v} < v_1; v) &= p_0(v)v - p_0(v_0)v_0, \\ U_B(v_1 \leq \hat{v} \leq \overline{v}; v) &= v - p_0(v_0)v_0 - p_1(v_1)v_1. \end{aligned}$$

It is easy to see that, when $v < v_0$:

$$U_B(v; v) = U_B(\underline{v} \leq \hat{v} < v_0; v) > U_B(v_0 \leq \hat{v} < v_1; v) > U_B(v_1 \leq \hat{v} \leq \overline{v}; v),$$

where the first inequality follows from assumption (A2) and the second follows from (A2) and the fact that p_1 is strictly increasing; when $v_0 \leq v < v_1$:

$$U_B(v; v) = U_B(v_0 \leq \hat{v} < v_1; v) \geq U_B(\underline{v} \leq \hat{v} < v_0; v) > U_B(v_1 \leq \hat{v} \leq \overline{v}; v),$$

where the first inequality follows from assumption (A2) and the second follows from (A2) and the fact that p_1 is strictly increasing; and when $v \geq v_1$:

$$U_B(v; v) = U_B(v_1 \leq \hat{v} \leq \overline{v}; v) \geq U_B(v_0 \leq \hat{v} < v_1; v) > U_B(\underline{v} \leq \hat{v} < v_0; v),$$

where the first inequality follows from the fact that p_1 is strictly increasing and the second follows from (A2). Hence, incentive compatibility is satisfied for all v. Finally, notice that also (3) is satisfied and is binding for the lowest type. This completes the proof. □

The optimal mechanism, beyond the null contract, involves a menu of two contracts: one in which the good is transferred for sure and a relatively high payment is required, intended for high types ($v \geq v_1$); one, intended for intermediate types ($v_0 \leq v < v_1$), in which the required payment is low and the allocation depends on the signal, but in a counterintuitive fashion: the allocation is $\pi^* = 1$ in case of bad signal ($s = 0$), it is zero otherwise. Therefore, the (interim) expected allocation, which is

$$E[\pi^*(v); v] = \begin{cases} 0 & \text{if } \underline{v} \leq v < v_0 \\ p_0(v) & \text{if } v_0 \leq v < v_1 \\ 1 & \text{if } v_1 \leq v \leq \overline{v} \end{cases},$$

is strictly decreasing in $[v_0, v_1)$. Hence, in this interval, there is inefficient rationing: higher types are assigned the good less often than lower types.

The reason why, for the seller, it is optimal to impose inefficient rationing is that this allows to extract surplus from relatively low types, and, at the same time, does not leave much information rent to high types. To see this, start considering a buyer of type v_0. This buyer knows that, if she reports truthfully, she is pretty likely to receive the good (because, for her, the signal is likely to be $s = 0$). If, instead, she underreports $\hat{v} < v_0$, she will certainly not receive the good. Hence, by reporting truthfully, this buyer would enjoy a relatively large gross surplus, that is extracted by the seller through a proper transfer.[7] Consider, now, type v_1 (or any type above): this type knows that, if she reports truthfully, she will receive the good for sure; if, instead, she mimics type v_0 (or any type $v \in [v_0, v_1)$), she is very likely not to receive it, because, being v_1 relatively high, the probability that the signal will be $s = 0$ is pretty low. In other words, by mimicking v_0, type v_1 will face the same (low) transfer as v_0, but not the same probability of getting the good, because the allocation, depending on the signal, eventually depends (negatively) on the *true* type. Hence, the fact that the allocation rule decreases with the (true) type when the report is in the interval $[v_0, v_1)$, has the effect of making it less attractive for high types to underreport their type, thereby allowing the seller to extract a large amount of surplus from them.

4 Applications

The previous literature highlighted that, when all the terms of the contract can be conditioned on an ex-post signal (correlated with the agent's type) and there are no restrictions on feasible payments, surplus maximization (i.e. efficiency) can be achieved and the principal can extract all the surplus from the agent. On the other hand, as outlined in the previous section, when only the allocation can depend on the signal, the optimal mechanism is clearly inefficient (types $v < v_1$ are assigned the good with probability less than one), and leaves some information rent to types $v > v_0$. Therefore, the impossibility of making payments contingent on the post-contractual signal reduces the principal's payoff. A natural question is then how realistic our restriction is.

Apart from situations in which it is explicitly prohibited, there may be more than one reason why the principal may prefer not to resort to contractual clauses that envisage future conditional payments by the agent. For example, in some cases, obtaining the agreed payment by the agent in the future may be difficult or very costly, whereas a modification in the allocation may be directly and immediately implemented by the principal. In others, delayed payments are not imposed because they cannot be enforced.

[7] Assumption (A2) guarantees that all types $v \in (v_0, v_1)$ do not find it convenient to underreport $\hat{v} < v_0$ either.

In this section, we present two applications in which, for some reason, the principal prefers not to require delayed contingent payments, but still uses the ex-post information to determine the terms of the agreement. We will show that these situations can be reinterpreted in light of our abstract model.

4.1 Internet traffic plans

The first application we consider is a real one and is represented by the internet traffic plans offered by wireless providers to mobile phone users. Operators typically offer plans with bandwidth caps, i.e. contracts of the type (K_i, t_i), where t_i is the price paid by the user to receive an endowment of K_i Gigabytes of internet traffic. If the user depletes her traffic endowment, then the connection speed is greatly reduced. These tariffs can be interpreted as contracts where the allocation (the connection speed) is contingent on an ex-post signal (the amount of internet traffic consumed). The price, instead, is paid upfront and no further payments are required.

To see this, consider an individual who needs internet access to perform a task through her mobile phone. For example, this task could be participating to a conference call, downloading a dataset or uploading a buy or sell order. Let v represent the utility that this individual obtains from completing the task; if she is not able to do so, her utility is equal to zero. Hence, the utility function of a type-v individual is $u(v) = \pi \times v - t$, where $\pi = 1$ $[\pi = 0]$ if the individual is able [unable] to perform the task and t is the price paid to get internet access. The individual does not know in advance how much internet traffic she will need to perform the task: in particular, she does not whether she will need more or less than K Gigabytes. Let p_1 be the (commonly known) probability that K Giga will not be enough for the individual and assume that p_1 is strictly increasing in v: more valuable tasks typically require more Giga to be completed.

The operator can easily monitor the internet usage made by the individual: therefore, the amount of internet actually consumed can be used as a signal of the true valuation she attaches to the completion of the task. If we assume that providing internet acces has no (marginal) cost for the wireless operator, this situation is isomorphic to our theoretical model. In particular, applied to this context, the optimal mechanism is the following menu of contracts:

$$[(K_0 = 0, t_0 = 0); (K_1 = K, t_1 = p_0(v_0) \cdot v_0); (K_2 = \infty, t_2 = p_0(v_0) \cdot v_0 + p_1(v_1) \cdot v_1)] .$$

Beyond the null contract (no internet access), the operator offers one contract with low price and K Gigabytes of high-speed traffic and one with high price and unlimited high-speed traffic. What was the allocation in the theoretical model is the supply of the (high-speed) traffic needed by the individual in any moment: the contract with threshold K is one in which the needed traffic is given to the individual (the allocation is equal to one) only so long as she has not reached K, which occurs with probability $p_0(v)$; as soon as the threshold is hit, high-speed traffic is interrupted (the allocation is equal to zero). In the contract with unlimited traffic, the threshold ($K_2 = \infty$) is never reached, so the allocation is always equal to one.

Some remarks: (i) strictly speaking, here the signal is perfectly correlated with the agent's actual demand of traffic; the latter, however, is imperfectly correlated with the agent's valuation: at the moment of signing the contract, the agent does not know the exact amount of traffic needed, but this is positively correlated with her valuation. As a result, the signal is imperfectly correlated with the agent's private information, just like in our abstract model; (ii) in our theoretical model, the signal is exogenous; however, in this application, the signal – the depletion of the internet traffic endowment – is potentially endogenous, hence manipulable. Notice, however, that, in the optimal mechanism, there is no advantage from manipulating the signal: once the contract has been chosen, the individual has nothing to gain by consuming more or less Gigabytes than those she actually needs; (iii) in the description above, to perfectly match our theoretical model, we assumed an exogenous threshold K; in reality, the threshold can be chosen by the operator which, moreover, can offer a multiplicity of contracts with different thresholds.

4.2 Promotion schemes

The second application considers a promotion scheme that can fit in hierarchical organizations where wages are exogenously set and fixed. In these institutions, usually public, an employee can climb the wage ladder only through promotions.[8] Our guess is that, in these organizations, there may be an implicit contract between employer and employee, whereby the employee, in order to be entitled for a promotion, must comply with some non-contractual obligations or activities.

To see this, consider an institution with two hierarchically ordered divisions. The agent is a worker currently employed in the low division where she is paid a fixed wage (normalized to zero). The employee would like to be promoted to the high division, where wages are, at least partially, related to the productivity v.[9]

Promotions are decided by the supervisor of the low division (the principal in the model). The supervisor's payoff is totally unaffected by the productivity of the worker

[8]For evidence that, also in the private sector, promotions play the major role in determining wage levels, see, e.g., Van Herpen et al. (2006).

[9]Alternatively, one can think of a situation where the wage in the high division is fixed as well, but a promotion will improve the reputation of the worker, allowing her to enjoy extra-benefits, these related to her true skills (e.g. consultancy, conference invitations, political appointments). For example, in the University system in the UK, senior lecturers and readers have the same wage level, but the latter is considered a more prestigious position as it entails positive reputation effects.

(the idea is that the low-level job does not require any particular skill, so every worker is equally able to do that); also, the supervisor is totally uninterested in the promotion decision: her primary interest is to make the division work properly (because, for example, she will receive bonus for that). To reach her goal, the supervisor may ask the worker to engage in additional (unpaid) activities that go beyond what is contractually required from her.

If we denote by π the promotion decision ($\pi = 1$ means promotion, $\pi = 0$ means no promotion), and by t the amount of extra-activities performed by the worker, the payoffs of the worker and the supervisor are, respectively,

$$U_W = \pi \times v - t, \qquad\qquad U_S = t,$$

where it is assumed that t also measures the disutility for the worker as well as the benefit for the supervisor associated with a level t of extra-activities.

By definition, the extra-activities cannot be required explicitly in a contract; however, there may be an implicit agreement between the two parties that they represent a necessary condition for the worker to get promoted.

Suppose, finally, that, after this implicit contract has been stipulated, the supervisor observes a binary signal s that is positively correlated with the true productivity of the worker, with $p_1(v)$ being the probability of observing a good signal ($s = 1$) when the worker's true productivity is v. In this case, the elements of the agreement, π and t, can, in principle, be made contingent on the signal itself. Notice, however, that, making the extra-activities contingent on the signal would probably not be a good idea for the supervisor: in fact, once the signal has realized and the promotion decision has been made, the worker will no longer have an incentive to comply with the agreed activities, and, being the contract an implicit one, there would be no way to enforce it. Therefore, it would probably be better for the supervisor not to make the extra-activities contingent on the signal, but requiring that these are performed immediately. Still, the supervisor can benefit from a contract that uses the ex-post signal to condition the promotion decision.

The situation outlined above perfectly matches the hypothesis behind our theoretical model, where the extra-activities play the same role of the payments in the buyer-seller setting.

Hence, applied to this context, the optimal mechanism corresponds to the following implicit agreement: (i) if no additional activity is performed, the worker will never be promoted in the future: this will be the choice made by low-productivity workers; (ii) the worker will be promoted for sure if she executes a high level of additional activities: high productivity workers will opt for this; (iii) if, instead, the worker performs only a modest level of extra-activities, she will be promoted only if the ex-post signal turns out to be bad: this will be the choice made by workers with intermediate productivity.

5 Effects on welfare

The optimal mechanism derived in Section 2 is clearly inefficient: buyers with valuation $v < v_1$ are assigned the good with probability less than one, whereas efficiency would prescribe the good to be transferred regardless of the buyer's type. Moreover, the buyer is rationed inefficiently: in the interval $[v_0, v_1)$, lower types are given the good more often than higher types.

One question that naturally arises is then whether the use of the ex-post signal in the determination of the allocation increases or decreases social welfare relative to a situation where no signal is available (or, even if available, it is not used).

Now, without the ex-post signal, the optimal mechanism is implemented by a single take-it-or-leave-it offer at price $\tilde{v}$, where $\tilde{v}$ is such that $\tilde{v} - (1 - G(\tilde{v}))/g(\tilde{v}) = 0$.[10] This offer is accepted only by types $v \geq \tilde{v}$, i.e. the expected allocation is

$$E[\pi^*(v); v] = \begin{cases} 0 & \text{if} \quad \underline{v} \leq v < \tilde{v} \\ 1 & \text{if} \quad \tilde{v} \leq v \leq \overline{v} \end{cases} .$$

If we denote by SS and by $\tilde{SS}$ the social surplus associated with the optimal mechanism with and without the ex-post signal, respectively, then their difference is given by

$$SS - \widetilde{SS} = \int_{v_0}^{\tilde{v}} v p_0(v) g(v) dv - \int_{\tilde{v}}^{v_1} v p_1(v) g(v) dv.$$

The first term in the expression above is the welfare gain associated to the fact that, with ex-post signal, types in $[v_0, \tilde{v})$ are assigned the good with strictly positive probability (whereas they would never receive it in the absence of the signal); the second term is the welfare loss associated to the fact that, with ex-post signal, types in $[\tilde{v}, v_1)$ are assigned the good with probability less than one (whereas they would receive it for sure in the absence of the signal).[11]

It turns our that the total effect on welfare is ambiguous: the social welfare may increase or decrease depending on the specific parameters of the model. To see this, consider the following example with uniform types and $p_1(v)$ linear.

Example. *Suppose that $v \sim U[0,1]$ and that $p_1(v) = kv$, with $k \leq 1$. Then: $\tilde{v} = 1/2$, $v_1 = 2/3$, $v_0 = (1 + k - \sqrt{k^2 - k + 1})/3k$, and $SS > \widetilde{SS}$ if and only if $k < \hat{k} \approx 0.71$.*

The example above suggests that the effect on welfare of the presence of the signal may be related to the symmetry of the problem. To see this, notice that a value of k close to

[10] This result is standard. It can also be obtained from our previous analysis by assuming that $p_0(v)$ is constant (i.e. the signal is totally uninformative).

[11] Notice that assumption (A1) ensures that $v_0 < \tilde{v} < v_1$. This also implies that, when the signal is available, the seller strictly prefers to use it.

one means that the signal is highly symmetric in the sense that the likelihood of observing a good signal from a high type is similar to the likelihood of observing a bad signal from a low type. The lower the value of k, the more asymmetric the signal is. Therefore, it seems that, when the problem at hand is highly symmetric, the use of the signal reduces social welfare. The following proposition shows that this intuition is correct.

PROPOSITION 2. *If, for all* $0 \leq \Delta \leq \min\{\tilde{v} - \underline{v}; \overline{v} - \tilde{v}\}$,

(i) $p_0(\tilde{v} - \Delta)g(\tilde{v} - \Delta) = p_1(\tilde{v} + \Delta)g(\tilde{v} + \Delta)$, *and*

(ii) $\frac{d}{dv}\left[-(1 - G(v - \Delta))(v - \Delta)p_0(v - \Delta)\right] = \left|\frac{d}{dv}\left[-(1 - G(v + \Delta))(v + \Delta)p_1(v + \Delta)\right]\right|$,

then $SS < \widetilde{SS}$.

Proof. Consider condition *(ii)*. If we divide the LHS by $p_0(\tilde{v} - \Delta)g(\tilde{v} - \Delta)$ and the RHS by $p_1(\tilde{v} + \Delta)g(\tilde{v} + \Delta)$, we obtain $\psi_0(\tilde{v} - \Delta) = |\psi_1(\tilde{v} + \Delta)|$, i.e. virtual valuations are symmetric with respect to $\tilde{v}$. This immediately implies that $v_1 - \tilde{v} = \tilde{v} - v_0$. This last fact, together with condition *(i)*, implies that

$$\int_{v_0}^{\tilde{v}} p_0(v)g(v)dv = \int_{\tilde{v}}^{v_1} p_1(v)g(v)dv.$$

We then have the following chain of inequalities:

$$\int_{v_0}^{\tilde{v}} vp_0(v)g(v)dv < \int_{v_0}^{\tilde{v}} \tilde{v}p_0(v)g(v)dv = \int_{\tilde{v}}^{v_1} \tilde{v}p_1(v)g(v)dv < \int_{\tilde{v}}^{v_1} vp_1(v)g(v)dv,$$

i.e.

$$SS - \widetilde{SS} = \int_{v_0}^{\tilde{v}} vp_0(v)g(v)dv - \int_{\tilde{v}}^{v_1} vp_1(v)g(v)dv < 0.$$

□

The first condition in Proposition 2 says that, broadly speaking, the probability that a buyer with a low valuation (relative to $\tilde{v}$) is drawn and a bad signal is observed is equal to the probability that a buyer with high valuation (relative to $\tilde{v}$) is drawn and a good signal is observed. The second condition is essentially a symmetry condition in the virtual valuations associated with the two signals. It can be given the following interpretation: suppose that the seller considers using a menu of three contracts with the same structure as the optimal one, with threshold values, v_0 and v_1, that have the same distance from $\tilde{v}$. The condition states that the marginal revenue for the seller associated with a small decrease in v_0 is equal to the marginal revenue associated with a small increase in v_1.

Notice that, when the distribution of valuations is uniform over [0,1], then the two conditions reduce to the single condition $p_0(\frac{1}{2} - \Delta) = p_1(\frac{1}{2} + \Delta)$, i.e. the signal is perfectly symmetric with respect to $\tilde{v} = 1/2$ (this corresponds to the case $k = 1$ in the example above).

6 Conclusion

We studied a contractual design problem between a seller and a buyer when some information is publicly observed ex-post and the allocation but not payments can be made contingent on it. Our analysis showed that, to increase her profit, the seller should include, in the menu of contracts offered to the buyer, one which the good is transferred only if the signal turns out to be bad. This generates inefficient rationing: some buyers with low valuation are assigned the good more often than others with higher valuation. We also showed that, relative to the case in which no signal is available (or, even if available, it is not used), the optimal contract may increase or decrease social welfare. However, when the problem at hand is sufficiently symmetric, the effect on welfare is certainly negative. This suggests that, when, for some reasons, contingent payments cannot be included, it may be preferable from an efficiency perspective to restrict the seller to offer contracts where the other contractual provisions (e.g., the allocation of the good) cannot be conditioned on delayed information. In this respect, the paper contributes to the literature showing that there may be in situations where imposing restrictions on the contracts that economic agents can privately write may enhance welfare (see, e.g., Aghion and Hermalin, 1990).

The model has been kept as simple as possible to focus on its main implications, but can be extended in several directions. The extension to an n-dimensional signal is straightforward but does not seem to offer new insights. Clearly, the number of contracts necessary to implement the optimal solution will increase with the number of outcomes. However, the main result that the expected allocation is non monotone in the buyer's type will remain valid. Relaxing assumption (A1) has very standard consequences on the equilibrium: types which choose different contracts in our equilibrium will bunch together.[12] Instead, the violation of assumption (A2) is fatal: no incentive-compatible contract can take advantage of the additional information and therefore the static lemon outcome will emerge.

[12]See Krishna (2010) for a standard treatment of bunching in direct revelation models.

References

[1] Aghion, P., Hermalin, B. (1990). Legal restrictions on private contracts can enhance efficiency. *Journal of Law, Economics, & Organization* 6, 381-409.

[2] Bose, S., Zhao, J. (2007). Optimal use of correlated information in mechanism design when full surplus extraction may be impossible. *Journal of Economic Theory* 135, 357-381.

[3] Cremer, J., McLean, R.P. (1988). Full extraction of the surplus in Bayesian and dominant strategy auctions. *Econometrica* 56, 1247-1257.

[4] Demougin, D.M., Garvie, D.A. (1991). Contractual design with correlated information under limited liability. *RAND Journal of Economics* 22, 477-489.

[5] Kosmopoulou, G., Williams, S. (1998). The robustness of the independent private value model in a Bayesian mechanism design. *Economic Theory* 12, 393421.

[6] Krishna, V. (2010). *Auction theory.* Academic press.

[7] McAfee, R.P., Reny, P.J. (1992). Correlated information and mechanism design. *Econometrica* 60, 395-421.

[8] Riordan, M.H., Sappington, D.E. (1988). Optimal contracts with public ex post information. *Journal of Economic Theory* 45, 189-199.

[9] Robert, J. (1991). Continuity in auction design. *Journal of Economic Theory* 55, 169-179.

[10] Van Herpen, M., Cools, K., Van Praag, M. (2006). Wage structure and the incentive effects of promotions. *Kyklos* 59, 441-459.