Communication Costs and Incentives to Acquire Soft and Hard Knowledge

Antonio Abatemarco*
University of Salerno & CELPE

Alberto Bennardo†
University of Salerno & CSEF

(preliminary draft)

Abstract

We study a multiple tasking principal-agent model where the agent gathers soft and hard knowledge for operational purposes. Within this set-up, we model communication from the agent to the principal as the process of hardening and transmitting soft knowledge, in the spirit of Dewatripont and Tirole (2005), and we assume that soft information, once hardened, can be used by the principal as a measure of the agent contribution to production (e.g., for incentive purposes). The assumption that hard and soft knowledge are complements in the communication technology, which naturally reflects the non-depletable nature of hard knowledge, leads to the following results. In the second best, the principal centralizes knowledge acquisition choices, and imposes the agent to acquire more knowledge than he would freely do, were his decision reflect market prices. Relatedly, if monitoring costs prevent the two parties from contracting the agent's knowledge acquisition activities (e.g., under full delegation), the agent underinvests in knowledge. These results are proved first within a set-up where hard knowledge is a homogeneous good and, subsequently, generalized by considering the case where several types of knowledge can be acquired, and gathering technologies differ with respect to the intensity in the use of non-monetary and monetary inputs.

Keywords: *hard knowledge, soft knowledge, communication, agency*
JEL: *D20, D82, D86*

*Department of Economics and Statistics and CELPE, University of Salerno, Via Giovanni Paolo II, 132 - 84084 - Fisciano (SA), Italy. Email: aabatemarco@unisa.it

†Department of Economics and Statistics, University of Salerno, and CSEF, University of Napoli "Federico II". Email: abennardo@unisa.it (corresponding author).

1 Introduction

Professional workers, managers, executives and white collars as well as entrepreneurs devote most of their working time to gather and exploit information and knowledge for operational purposes. Roughly, these activities can be distinguished into two subsets: acquisition of hard or explicit, transferable knowledge and information, and acquisition of specialized, soft knowledge which resides with the agent, and is difficult to express, formalize and substantiate into transferable signals. Managerial competence, technical and scientific knowledge, or rough data on the economic fundamentals of the environment in which an agent operates, are all common examples of hard information, whereas knowledge obtained by addressing specific problems through trials, learning by doing process, or, informal, not fully articulated, reasoning guided by past experience are examples of soft knowledge.

The key distinction between hard and soft knowledge, which goes back to Hayek (1945) and Polanyi (1966), reflects their codifiability, namely the extent to which knowledge can be formally expressed, classified and coded (see Garicano and Wu (2012) among others): completely soft (tacit) knowledge cannot be codified; at the other extreme, hard knowledge can be perfectly communicated by using standard codes, which are infrastructures that can be repeatedly operated once created ("communication channels" in the terminology coined by Arrow (1974)).

A large literature starting from Crawford and Sobel (1982) studies whether and how agents can transmit soft information; their approach has more recently been adopted in the organization literature (see, for instance, Dessein 2002). Nonetheless, several authors including Hayek (1945), Dewatripont and Tirole (2005), Liberti et al. (2018), Stein (2002) have stressed that most information is neither completely soft nor perfectly hard, "it is in the middle", and, most importantly, that the degree of softness of soft knowledge is endogenous: soft information can be, at least partially, hardened, by exerting costly communication activities. For example, a formal proof of a math theorem hardens a mathematical conjecture, an information sharing system across banks hardens the knowledge on borrowers indebtedness etc.

This paper makes a step toward being more explicit in modeling the communication technology that agents' use to harden soft information. A key feature of our analysis is that (previously gathered) hard knowledge is the main asset (infrastructure) that agents use to communicate, i.e. to harden soft knowledge. We explicitly introduce a communication technology with this property within a Principal-Agent multiple-task setting, where the effort exerted by the agent in gathering and using soft information is unobservable to the principal (due to the essential inarticulate nature of soft knowledge), whereas costly activities undertaken to acquire hard knowledge can be made verifiable by the principal, possibly at a positive cost. Within this set-up, we analyze the agent's private incentives to acquire hard knowledge and to communicate soft knowledge, as well as the net gains the principal can achieve by delegating (or centralizing) hard knowledge gathering choices to the agent.

Our model sheds some new lights on these organizational and contractual choices, by showing that a natural complementarity between hard and soft knowledge — which is featured by the communication technology and reflects the non-depletable nature of hard knowledge — leads to the following results. Under full delegation of information gathering choices, the agent private incentives to gather hard information fall short of social incentives; consequently, in the second best the principal imposes the agent to gather more hard information than he would freely do were his decision reflect market prices (e.g., under full delegation).

As a matter of fact, the trade-offs underlying these choices shape the organization and the contractual structures of many real life institutions. Most large companies are organized into business units, but differ in the degree of centralization of activities such as product development, accounting and finance that largely rest upon the acquisition and the use of hard knowledge. Similarly, network of firms setting contractual agreement aimed at developing scientific collaboration or information sharing infrastructures, differs in the degree of centralization of hard knowledge and information gathering activities. By the same token, a fundamental difference between standard credit contracts and venture capital credit agreements is that in contrast to banks, contracts between venture capitalists (the principals) and agents typically involve the venture capitalists centralizing important decision concerning the professionalization of the management (the acquisition by the firm of formal knowledge).

Throughout the paper, we shall postulate that hardening soft knowledge requires learning and codification activities, encompassing both the acquisition and the use of explicit knowledge. In other words, explicit knowledge is the key input — the infrastructure — that economic agents use to harden soft knowledge. This is for two reasons.

First, in order to communicate his tacit knowledge an agent needs a codified language, a set of classification protocols and conventions, as well as the formal knowledge embedded in the technology (softwares) he uses to communicate. For instance, the director of a local branch of a bank must learn how to use a grading rules for loans and borrowers (communication protocols), which are often standardized through a proprietary software, in order to pass local information about the riskiness of his clientele through the company's hierarchy.

Second, in many real life situations the only way an agent can transmit the tacit knowledge that he has obtained through learning by doing, or through intuitive judgements based on past experience, is to acquire the cognitive ability to make "formal statements" replacing his heuristics. For this purpose, the agent needs to learn how to use formal knowledge produced by other agents (theorems, formal logic, causal analysis etc.) to develop this knowledge by himself. For instance, the branch director of a bank or a division manager may have observed some local information (say a signal on the state of demand) that has induced him to update his prior beliefs. Then, knowledge of elementary principles of Bayesian inference enables him to communicate why and how he updated his prior distribution, and to replace somewhat confusing informal statements about his beliefs' change. Similarly, knowing some econometrics may enable him to provide better evidence and to argue more forcefully on the existence of a causality relationships between observable variables, that he has understood through long periods of field experience.

Within our model, communication activities take place in a multiple tasking set-up where a risk-neutral principal and an agent, whose risk attitude is conventionally described by CARA preferences, interact through several stages. At the initial stage, the principal offers to the agent an exclusive contract that the latter can either accept or refuse; subsequently the agent engages in two costly activities, acquisition and use of specialized human capital (specific knowledge) and acquisition of hard knowledge, which is used subsequently to better evaluate a signal on the state of the world. In an intermediary stage, the agent observes two imperfect signals: one provides information on the true state of nature (for instance a demand or a supply shock that is realized in the final stage); the other signal, which in our setting can be naturally interpreted as an intermediary output (such as, for instance, a prototype resulting from research activity). The quality of the signal on the true state of nature depends on the amount of hard knowledge gathered by the agent, whereas the expected value of the intermediary output is increasing in the amount of specific knowledge.

The agent then uses a communication technology to transmit his information on the value of the intermediary output and on the true state of nature to the principal. The signal on the state of nature is hard information and can be perfectly transmitted, whereas the value of the intermediary output is soft information that needs to be hardened. The hardening process produces a signal whose precision is increasing both in the amount of hard and soft knowledge gathered by the agent. We assume, in the spirit of the literature on certifiable information (Milgrom 1981), that all knowledge that is hardened by the agent is truthfully communicated to the principal, whereas information that remains tacit is not contractually used in the agency relationship.

Finally, the information revealed at the intermediary stage (the signal on the state of nature and the signal on the intermediary output communicated by the agent) is used by the principal in the last stage to choose a monetary investment (for operational purposes), production is realized and the agent receives a payment contingent on the realization of the uncertainty resolved at the intermediary and at the final stage, as well as on the information he communicates.

Within this set-up, the paper addresses the following two questions. First, under which conditions is it second best optimal for the principal to let the agent "freely" choose how much hard information to acquire? Second, whenever full delegation is not constraint efficient, do second best arrangements impose the agents to gather more or less hard information than he would freely do under a contract that provides only monetary incentives?

The following assumption of complementarity between soft and hard knowledge in the communication process is imposed in characterizing the second best contract of this problem. The larger the amount of specific tacit knowledge gathered by the agent, the stronger the impact of explicit hard knowledge in reducing the noise in the transmission of information by the agent to the principal. We argue that the very essential feature of formal knowledge, namely is its non depletability, generates this complementarity. An agent involved in his core production activity for along time acquires several pieces of soft information, and can use the same formal knowledge (the same "information channel" in the language of Arrow) repeatedly, in order to communicate each of these pieces. More generally, when the amount of specialized knowledge gathered on a specific problem gets larger, *ceteris paribus*, the expected value that an additional "bite" of formal knowledge yields in the communication process increases as well, because of the non depletability of formal knowledge.

Under this assumption, our analysis delivers the following results. In the second best the principal imposes the agent to acquire more hard knowledge than he would freely choose. Consistently, under the second best contract, the shadow price of hard knowledge paid by the agent is lower than the market price. We also study the case in which acquisition of hard information cannot be verified, and show that, under mild convexity assumptions, it is lower than the corresponding second-best level.

Finally, we extend these results to a set-up where the agent can acquire multiple types of knowledge, gathering knowledge requires both money and effort, and knowledge can be either *supervisory* or *executive* in the sense of Dewatripont and Tirole (2005). We show that imposing the agent to acquire more knowledge of a certain type, say type j, than he would freely do is second best optimal if the magnitude of the complementarity between soft and hard knowledge in the communication technology overcomes the substitution effect between different types of effort (exerted in production and in information gathering) traditionally highlighted by the multiple tasking literature.

By acquiring more information than the amount that maximizes his expected utility — given the contractual payments — the agent reduces the variability of signals' realizations and, hence, the variability of his consumption. This relaxes his incentive constraint with respect to soft knowledge acquisition, e.g., it makes it relatively more valuable to exert effort in that activity. In choosing hard knowledge under full delegation, the agent does not internalize the effect of his choice on this incentive constraint, and, hence, on the amount of insurance that he can obtain from the principal.

These results provide a purely incentive based rationale for why economic institutions, in the presence of endogenous communication costs limiting the transfers of soft information, centralize the functions involving the acquisition of hard information. The same findings suggest that contractual arrangements through which a venture capitalist providing funds "imposes" an entrepreneur to acquire managerial skills and professionalization. This practice can be rationalized as second best devise that the venture capitalist (the principal) uses in order to facilitate the evaluation of the entrepreneurial activities at intermediary stages and the assessment of firms' investment opportunities.

2 The Model

We consider a multi-task principal-agent set-up in which a risk-averse agent, after signing a contract with a risk-neutral principal, operates a two-stages production technology. In the first stage, he exerts effort to use and develop his specialized knowledge (his idea, project, or specific human capital), and gathers general knowledge (managerial competence, professionalization, or technical and scientific knowledge) as production inputs. General knowledge can be also acquired by the principal through outsourcing, whereas specialized knowledge is exclusively provided by the agent.

These two types of knowledge are then used in the second stage. Specialized knowledge produces an intermediary output, whereas general knowledge provided by the agent has a double role: it allows the agent to communicate (to harden) the soft knowledge that he has gathered as well as to better interpret hard information (information on the state of nature).

In the following we describe first the production and then the communication technology.

Production and information acquisition

Let $a \in A$ and $z_A \in Z$, with A and Z being compact sets, be the effort exerted by the the agent in gathering specialized knowledge and the amount of general knowledge that he acquires in the first stage, respectively; a is assumed to generate a disutility cost $\varphi(a)$ while z is acquired at some monetary cost q_A.

The action a produces an intermediate output $x = h(a)$, where $h(\cdot)$ is a strictly concave, increasing and twice differentiable function of effort.

The intermediary output x together with a monetary investment $m \in M = [0, +\infty)$ produces a final output, $y = g_1(x + \omega_1) + g_2(m + \omega_2)$, whose value depends on the realization of two states of nature, $\omega_1 \in \Omega_1$ and $\omega_2 \in$

Ω_2, affecting the marginal productivity of x and m respectively; both ω_1 and ω_2 are normally and independently distributed with means μ_{ω_1} and μ_{ω_2}, and variances $\sigma^1_{\omega_1}$ and $\sigma^2_{\omega_2}$ respectively. We will assume that both $g_1(\cdot,\cdot)$ and $g_2(\cdot,\cdot)$ are strictly concave and twice differentiable in $(x+\omega_1)$ and $(m+\omega_2)$ respectively.

At the end of the production process, the principal and the agent observe both $(x+\omega_1)$ and ω_2. However, the true state of nature ω_2 is unknown to both the agent and the principal at the second production stage (e.g., when m is chosen). At that stage, the principal can use the general knowledge acquired by the agent, z_A, as well as that obtained through outsourcing, z_P, to interpret a signal $\theta \in \Theta$ of the true state ω_2. We assume that θ is normally distributed with mean μ_{ω_2} and variance $\sigma^2_\theta(z)$, where $z = (z_A, z_P)$ is the overall amount of general knowledge acquired by the agent and by the principal. The larger is z_A the better the communication ability of the agent, thereby the lower the noisiness of the signal the agent transfers to the principal.

Information and communication

The effort a exerted in using and gathering specialized human capital is assumed to be private information of the agent and it is not contractible. In the first part of the paper, we shall assume that z_A is verifiable at zero cost and contractible. Subsequently, we shall also study an alternative regime where z_A is unobservable and, for this reason, the choice of z_A must be fully delegated to the agent.

The signal θ is observed by the principal, whereas the intermediary output x is soft information observed by the agent only. The information about the value of the intermediary output, x, can be communicated by the agent to the principal only to the extent that it is hardened by the agent. The outcome of the communication process is a signal x_c whose precision depends upon both the amount of soft and hard information gathered by the agent. More specifically, the signal $x_c = h(a) + \varepsilon_c$, where ε_c is a normally distributed error with zero mean and variance $\sigma^2_c(a, z_A)$.

The following two assumptions on communication are key for our analysis. First, communication skills of the agent are increasing in his knowledge z_A, so that, for all $a \in A$ and $z'_A, z_A \in Z$ with $z'_A > z_A$, $\sigma^2_c(a, z'_A) < \sigma^2_c(a, z_A)$.

Second, for all $z_A \in Z$ and $a', a \in A$ with $a' > a$, $\partial\sigma^2_c(a', z_A)/\partial z_A < \partial\sigma^2_c(a, z_A)/\partial z_A$, that is, the larger the amount of specific tacit knowledge gathered by the agent, the stronger the impact of explicit hard knowledge in reducing the noise in the transmission of information by the agent to the principal. It is natural to interpret a as a vector of many activities giving to the the agent 'pieces' of information to be communicated to the principal by using knowledge z_A which acts as a public good, in that the use of knowledge to communicate *specialized* human capital does not deplete its availability for further communication.

Contracts

Contractual payments can be made contingent on information that both the agent and the principal can observe and verify: namely x_c, θ, $(x+\omega_1)$, $(m+\omega_2)$, and y. However, as $(x+\omega_1)$ and $(m+\omega_2)$ are observed by both parties, y turns out to be a redundant signal.

As conventional in the literature, we shall restrict attention to linear contracts. A contract $\gamma := \{a, z_A, z_P, t, s_c, s_{\omega_1}, s_\theta, s_{\omega_2}\}$ then specifies a vector of actions (a, z_A, z_P), a fixed payment t, and a vector of linear (unitary) payments $s = (s_c, s_{\omega_1}, s_\theta, s_{\omega_2})$ contingent on signals $x_c, (x+\omega_1), \theta, (m+\omega_2)$ respectively.

Expected utility and profit

The agent is assumed to be risk-averse with constant absolute risk aversion (CARA), and preferences represented by the Bernoully utility

$$U(w(\gamma), a, z_A) = -e^{[-\eta(w(\gamma)-\varphi(a)-q_A z_A))]} \tag{1}$$

where $\eta > 0$ is the coefficient of absolute risk aversion, $\varphi(\cdot)$ is increasing and strictly convex, and $w(\gamma) = t + s_c x_c + s_{\omega_1}(x+\omega_1) + s_\theta\theta + s_{\omega_2}(m+\omega_2)$ is the linear payment function defined by the principal.

The expected utility of an agent who signs the contract γ and takes actions prescribed by the contract is

$$E[U(\gamma)] = E[-e^{-\eta(t+s_c x_c+s_{\omega_1}(x+\omega_1)+s_\theta\theta+s_{\omega_2}(m+\omega_2)-\varphi(a)-q_A z_A)}] \tag{2}$$

We do not restrict the sign of the correlation between x_c and ω_1, and between θ and ω_2;[1] all other pairs of primitive random variables are independently distributed. Hence, the certainty equivalent of $EU(\gamma)$ can be written as

$$\begin{aligned} CE(\gamma) = t + h(a)(s_c + s_{\omega_1}) + s_{\omega_1}\mu_{\omega_1} + (s_\theta + s_{\omega_2})\mu_{\omega_2} - \varphi(a) - q_A z_A + \\ -\frac{1}{2}\eta\left(s_c^2\sigma_c^2(a, z_A) + s_{\omega_1}^2\sigma_{\omega_1}^2 + 2s_c s_{\omega_1}\sigma_{c\omega_1}(a, z_A) + \right. \\ \left. + s_\theta^2\sigma_\theta^2(z) + s_{\omega_2}^2\sigma_{\omega_2}^2 + 2s_\theta s_{\omega_2}\sigma_{\theta\omega_2}(z)\right) \end{aligned} \tag{3}$$

Let the cost of the money investment be m, the certainty profit function of the principal is

$$\Pi(y, w, m, z_P) = y - w - m - q_P z_P \tag{4}$$

[1] Albeit it might be natural to assume that θ and ω_2 are positively correlated.

where q_P is the market price of general knowledge. The principal expected profit results from a two-stage maximization. In the second, last stage the money investment m is optimally chosen after observing x_c and θ. Hence, as m affects the function $g_2(\cdot)$ but not $g_1(\cdot)$, at that stage the principal chooses m by solving the following program.

$$m^*(\theta, z) = \arg\max_m \quad \int_{-\infty}^{+\infty} (g_2(m + \omega_2) - m)\, f(\omega_2|\theta, z) d\omega_2 \tag{5}$$

where $z = (z_A, z_P)$, and $f(\omega_2|\theta, z)$ is the *posterior* density function associated to $(m + \omega_2)$.

Provided that the agent chooses the actions prescribed by the contract γ, the expected profit that the principal obtains at the initial stage is then

$$\begin{aligned} E[\Pi(\gamma)] = &\int_{-\infty}^{+\infty} \int_{0}^{+\infty} g_1(x + \omega_1) f(\omega_1|x_c, a, z_A) f(x_c|a, z_A) d\omega_1 dx_c + \\ + &\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (g_2(m^*(\theta, z) + \omega_2) - m^*(\theta, z))\, f(\omega_2|\theta, z) f(\theta|z) d\omega_2 d\theta + \\ &- (t + h(a)(s_c + s_{\omega_1}) + s_{\omega_1}\mu_{\omega_1} + (s_\theta + s_{\omega_2})\mu_{\omega_2}) - q_P z_P \end{aligned} \tag{6}$$

In the first stage, the principal solves the second-best program, that is, he maximizes his expected profits under the incentive compatibility constraint (imposing that the agent acquires unobservable soft knowledge as prescribed by the contract), and the participation constraint:

$$\begin{aligned} \max_\gamma \quad & E[\Pi(\gamma)] \\ s.t. \quad & a = arg\max_a \quad E[U(\gamma)] \qquad (IC_a) \\ & E[U(\gamma)] \geq \bar{U} \qquad\qquad\qquad (PC) \end{aligned} \tag{7}$$

Later on in the paper, we shall also consider an alternative regime, hereafter full-delegation analysis, where the amount of hard knowledge gathered by the agent, z_A, is not observable by the principal, so that the decision on the acquisition of z_A must be fully delegated to the agent. Hence, the full-delegation contract solves program (7) under the following additional constraint

$$z_A = arg\max_{z_A} \quad E[U(\gamma')] \qquad\qquad (IC_z) \tag{8}$$

where the condition (IC_z) imposes that it is rational for the agent to acquire the amount of hard knowledge recommended by the program. In what follows, we indicate by $\gamma^{FD} = (a^{FD}, z^{FD}, t^{FD}, s^{FD})$ the solutions to the full-delegation program (8), whereas $\gamma^{SB} = (a^{SB}, z^{SB}, t^{SB}, s^{SB})$ are the solutions of the second-best program (7).

3 Second-best contract

We start the analysis of the second-best contract by highlighting a few properties that can be derived by using a standard concavity argument as follows, almost directly, from the first-order conditions of program (7).

To begin with, risk aversion implies that s_θ and s_{ω_2} are both equal to zero at the second best. Intuitively, setting $(s_\theta, s_{\omega_2}) = 0$ is (constrained) optimal as it minimizes the riskiness of the compensation without yielding any incentive loss with respect to any $(s_\theta, s_{\omega_2}) \neq 0$, as the bonuses s_θ and s_{ω_2} do not affect the incentive constraint.

Moreover, whenever the correlation between the two signals of the hidden action, x_c and $(x + \omega_1)$, is not too large, both the unitary payments s_c and s_{ω_1} are strictly positive at the second best. The intuition behind this result is quite standard (see Bolton and Dewatripont (2005)). First of all, in the spirit of Holmstrom's (1979) sufficient statistic theorem, the principal finds it optimal to use both signals of a (i.e., x_c and ω_1) whenever none of them is a sufficient statistic of the other. Second, whenever the correlation is negative or positive but not too large, optimal risk diversification for the agent commands s_c and s_{ω_1} both strictly positive at second best. This is essentially for the the same reason why an agent trading two assets in the CARA-Normal model buys positive amounts of both of them, whenever the correlation between these two assets is not too large.[2]

Finally, let $E[R(\gamma)]$, λ_1 and λ_2 indicate the expected revenue function (that is, the expected production) of the principal, the lagrange multipliers associated to the incentive compatibility and participation constraint respectively. Hence, the first order conditions with respect to z_A and z_P are, respectively,

$$\frac{\partial E[R(\gamma)]}{\partial z_A} + \lambda_1 \left(-\frac{1}{2}\eta s_c^2 \frac{\partial^2 \sigma_c^2(a, z_A)}{\partial a \partial z_A} \right) + \\ + \lambda_2 \left(-q_A - \frac{1}{2}\eta \left(s_c^2 \frac{\partial \sigma_c^2(a, z_A)}{\partial z_A} + s_\theta^2 \frac{\partial \sigma_\theta^2(z)}{\partial z_A} + 2 s_\theta s_{\omega_2} \frac{\partial \sigma_{\theta\omega_2}(z)}{\partial z_A} \right) \right) = 0 \quad (9)$$

and,

$$\frac{\partial E[R(\gamma)]}{\partial z_P} - q_P + \lambda_2 \left(-\frac{1}{2}\eta \left(s_\theta^2 \frac{\partial \sigma_\theta^2(z)}{\partial z_P} + 2 s_\theta s_{\omega_2} \frac{\partial \sigma_{\theta\omega_2}(z)}{\partial z_P} \right) \right) = 0 \quad (10)$$

As proven formally in the proof of Proposition 3.1 (see the appendix), the second-best contract sets $z_P = 0$ whenever the cost of acquiring knowledge for the principal is not much smaller than the cost of acquiring knowledge

[2] When the correlation coefficient is close enough to 1, setting s_{ω_1} or s_c negative is optimal for the same reason why an agent trading two assets featuring a large and positive correlation may find it optimal to sell short one of them in order to reduce the riskiness of his portfolio.

for the agent. Intuitively, this is because (i) z_A relaxes the agent incentive constraint as it makes it less uncertain the returns of action a by reducing the variability of the signal, whereas z_P has no effect on incentives, and (ii) an increase of z_A reduces the variability of the signal more than z_P.

Let $\rho_{c\omega_1}$ denote the correlation coefficient between x_c and ω_1. The following Proposition summarizes the discussion above.

Proposition 3.1. *The second-best contract satisfies the following properties: (I) $(s_\theta, s_{\omega_2}) = 0$; (II) there is $\bar{\rho}_{c\omega_1} \in (0, 1)$ such that $s_c > 0$ and $s_{\omega_1} > 0$ for any $\rho_{c\omega_1} \leq \bar{\rho}$; (III) there is $\Delta > 0$ such that $z_P = 0$ for any pair (q_A, q_P) such that $q_A \leq q_P + \Delta$.*

The proof of Proposition 3.1 as well as subsequent proofs are all provided in the appendix.

3.1 Second-best knowledge and decentralization

This section addresses the following two questions. First, under which conditions is it second best optimal for the principal to let the agent "freely" choose how much hard information to acquire ? Second, whenever full delegation is not constraint efficient, do second-best arrangements impose the agents to gather more or less hard information than he would freely do under a contract that provides only monetary incentives ? Let

$$z_A^* = \arg\max_{z_A} E[U(a^{SB}, z_A, z_P^{SB}, s^{SB})], \tag{11}$$

addressing the former of the two questions above amounts to ask whether the second-best value z_A^{SB} is different from z_A^*, or equivalently, whether the agent and the principal private incentives are aligned with respect to the choice of z_A. Addressing the latter of the two questions above amounts to verify whether $z_A^* \gtreqless z_A^{SB}$.

Next proposition shows that the second best incentives of the principal and the agent diverge with respect to the choice of z_A, implying that it would be suboptimal for the principal (at least whenever z_A can be monitored at a sufficiently low cost) to delegate the choice of z_A to the agent. Moreover, we will show that $z_A^* < z_A^{SB}$, which is the level of z_A that the agent would prefer to gather, were he free to choose, falls short of the second-best value.

Proposition 3.2. *Assume $(a^{SB}, z_A^{SB} + z_P^{SB}) > 0$, $z_A^{SB} > z_A^*$.*

The intuition for this result is as follows. First, starting from $z_A = z_A^*$, an increase of z_A allows the agent to provide a less noisy signal of the intermediary output. This reduces the variability of the agent's consumption and increases his utility. At $z_A = z_A^*$, however, this effect is exactly offset by the cost of z_A. Thus, the effect of a small increase in z_A around z_A^* on the agent utility is zero at first order. Moreover, an increase in the knowledge acquired by the agent has a non-negative direct effect on the principal's profit, roughly because the principal does not pay for the knowledge acquired by the agent. Finally, and most importantly, an increase in z_A makes more precise the signal of the agent's action, x_c. As a consequence, the agent's marginal gain from increasing a gets larger, or, equivalently, the incentive constraint gets relaxed. This allows the principal to reduce the variability of the agent's consumption by offering him a contract featuring less powerful incentives, and, ultimately, to get a higher profit. This explains why the principal gains from imposing the agent to gather more information than he would freely do at the second best: $z_A^{SB} > z_A^*$.

3.2 Second-best versus full-delegation knowledge

The analysis of the previous section has been developed under the assumption of contractibility of z_A.

In most real life situations, however, the knowledge gathered by an agent can be verified only when the competencies he obtained by acquiring that knowledge can be certified by a third party at a sufficiently low cost. Moreover, the real life cost of certification depends both upon the type of knowledge required to the agent (for instance, the acquisition of scientific knowledge is, in general, easier to certify than the acquisition of knowledge on the specific environment where a firm operates) and on the enforcement standards that prevail in the specific institutional context at stake.

For this reason, it becomes of interest, especially from a positive viewpoint, to understand whether the agent gathers more hard knowledge in the case where this knowledge is fully certifiable (e.g., in the second best) or when it is his own private information, respectively (e.g., in the full-delegation program).

Next proposition shows that the agent gathers more information in the second best program than in the full-delegation program.

Proposition 3.3. *Assume* $(a^{SB}, z_A^{SB} + z_P^{SB}) > 0$ *and* $\rho_{c\omega_1} = 0$, *the second best value of hard knowledge gathered by the agent,* z_A^{SB}, *is lower than the value of knowledge gathered under full delegation,* z_A^{FD}.

This proposition shows that starting from the full delegation contract, one can always find a contract that satisfies both the participation and the incentive constraint w.r.t. a, but not the incentive constraint w.r.t. z_A, and increases the expected profit of the principal. The intuition behind this result is as follows.

First, an increase in z_A of size dz_A, by reducing the variance of agent's consumption, would relax the incentive constraint w.r.t. a (i.e. would induce the agent to choose a larger a). The principal can then reduce the unitary bonus s_ω, by an amount ds_{ω_1}, and keep incentives to exert a unchanged.

Second, the principal can also vary t by dt so that (ds_{ω_1}, dt) leaves also the expected utility unchanged.

Then, it becomes key to ascertain how the perturbation $(dz_A, dt, ds_{\omega_1})$ affects the expected profit of the principal; as a matter of fact, dz_A increases the principal profit since z_A has positive operation value. Moreover, as we show in the proof, the perturbation (dt, ds_{ω_1}), that leaves the utility of the agent unaltered, increases the principal's expected profit. This is because the marginal rates of substitution between ds_{ω_1} and dt are different for the principal and the agent. Specifically, the latter values relatively more a unitary increase in t than a unitary increase in s_{ω_1}, as s_{ω_1} also increases the variability of the consumption.

4 An extension to multiple types of knowledge

In the baseline model developed in the previous sections, we treated formal knowledge as a homogenous good.

This section generalizes the analysis of the agent's incentives to acquire knowledge along the following two dimensions. We assume that the agent can acquire several types of knowledge and that a subset of types of knowledge can be acquired at a positive monetary cost only, whereas another subset requires both effort and money to be gathered (e.g. these types of knowledge also need effort to be understood). According to the distinction introduced by Dewatripont and Tirole (2005), it is possible to interpret knowledge of the former type as "supervisory" knowledge, which is used solely to identify the most appropriate course of action (for instance, knowledge that the agent can acquire from a consultant). This knowledge requires a negligible amount of effort to be used. Differently, the latter type of knowledge is "executive" in the language of Dewatripont and Tirole (2005): the agent must effectively fully understand ("incorporate") executive knowledge to take advantage of it, and for this reason acquiring this knowledge requires a non negligible amount of effort.

Introducing heterogeneity of knowledge broadens the scope of our analysis for the following reasons.

First, as discussed in the introduction, in the real world one can distinguish at least two different types of hard knowledge: (i) *coding* knowledge that agents utilize to communicate (e.g. to harden their soft knowledge); (ii) *scientific and professional* knowledge, that the agent uses to "replace" soft knowledge (heuristics). The highly stylized model analyzed in the previous sections did not distinguish between these two types of knowledge. Second, in the baseline model we neglected the possibility that acquiring knowledge requires effort having a negative impact on the marginal disutility of effort.

Formally, throughout this section we shall assume that the agent and the principal can acquire J different types of knowledge, and denote z_{jA} and z_{jP}, respectively, the amount of knowledge of type j acquired by the agent and the principal. Besides, we shall distinguish two disjoint subsets of J, J' and J'' with $J' \cup J'' = J$. The effort needed to understand any type of knowledge in the set J' is negligible, for simplicity equal to zero, whereas to acquire knowledge of type J'', the agent needs to exert a non negligible amount of effort. In particular, we shall assume, merely for simplicity, that for each 'unit' of knowledge z_j, the agent needs to exert $e_j = e(z_{jA})$ units of effort to understand it.

As in the baseline model, each signal serves two purposes in the principal-agent relationship: it enables the principal to better choose the monetary investment m, and it reduces the variance of the signal x_c.

Let a denote the action the agent devotes to the production of the final output y as in the baseline model, and e_j the effort he exerts to gather executive knowledge of type j, with j in J''.

We now model the disutility of effort as $\varphi(a + \sum_{j=1}^{J} e_j)$; the latter is the total cost the agent bears within the production process by supplying both specialized knowledge a and several types of general knowledge z_j.

Let $\mathbf{z_A} = (z_{1A}, ...z_{JA})$ and $\mathbf{z_P} = (z_{1P}, ...z_{JP})$ be the vectors of knowledge of different types acquired by the agent and the principal respectively.

The following separability assumption is introduced to simplify the analysis of the second-best contract.

Assumption 1. *$\sigma_c^2(a, \mathbf{z_A}) = \bar{\sigma}_c^2 - \sum_{j=1}^{J} \phi_j(a, z_{jA})$ with $\frac{\partial \sigma_c^2(a,\mathbf{z_A})}{\partial z_{jA}} < 0$, $\frac{\partial \sigma_c^2(a,\mathbf{z_A})}{\partial a} < 0$, and $\frac{\partial \sigma_c^2(a,\mathbf{z_A})}{\partial a \partial z_{jA}} < 0 \ \forall\ j$.*[3]

Let

$$\gamma^{\mathbf{SB}} = (a^{SB}, \mathbf{z_A^{SB}}, \mathbf{z_P^{SB}}, t^{SB}, s^{SB})$$

denote the second-best contract, and let

$$z_{jA}^{*} = \arg\max_{z_{jA}} \quad E[U(a^{SB}, \mathbf{z_{A-j}^{SB}}, z_{jA}, \mathbf{z_P^{SB}}, t^{SB}, s^{SB})]$$

be the amount of type j knowledge that maximizes the agent's expected utility at second-best contract.

Next proposition shows that at the second best the principal imposes the agent to overinvest in z_{jA} if and only if the complementarity between soft knowledge and hard knowledge in communication has a larger impact on incentives than the substitutability between effort needed to acquire knowledge and the productive effort a. In the opposite case the principal imposes under-investment.

[3]The assumption of separability is introduced to make the analysis as transparent as possible, but it can be easily relaxed by assuming, for instance, supermodularity of the function $\phi(\cdot, \cdot)$.

Proposition 4.1. *Assume* $(a^{SB}, z_{jA}^{SB} + z_{jP}^{SB}) > 0 \; \forall \; j \in J$*, the second-best vector of signals,* $\mathbf{z} = (\mathbf{z_A}, \mathbf{z_P})$*, features the following properties:*

I. for all $j \in J'$, $z_{jA}^{*} < z_{jA}^{SB}$*;*

II. for all $j \in J''$ *such that* $\frac{\partial^2 \sigma^2(a,\mathbf{z_A})}{\partial a \partial z_{jA}} + \frac{\partial^2 \varphi^2(a,e_j)}{\partial a \partial e_j} \frac{\partial e_j(z_{jA})}{\partial z_{jA}} < 0$ *for any possible vector* (a, z_{jA})*, then* $z_{jA}^{*} < z_{jA}^{SB}$ *whenever* $q_{jA} \leq q_{jP} + \Delta$*, with* $\Delta > 0$ *and small enough;*

III. for all $j \in J''$ *such that* $\frac{\partial^2 \sigma^2(a,\mathbf{z_A})}{\partial a \partial z_{jA}} + \frac{\partial^2 \varphi^2(a,e_j(z_{jA}))}{\partial a \partial e_j} \frac{\partial e_j(z_{jA})}{\partial z_{jA}} > 0$ *for any possible vector* (a, z_{jA})*, there is* $\Delta > 0$ *such that (i) if* $q_{jP} - q_{jA} < \Delta$ *then* $z_{jA}^{SB} = 0$ *and* $z_{jP}^{SB} > 0$ *for* q_P *small enough; whereas if* $q_{jP} - q_{jA} > \Delta$*, then* $z_{jA}^{*} > z_{jA}^{SB}$*, with* $z_{jA}^{SB} > 0$ *for* q_{jA} *small enough.*

Intuitively, under our assumption on preferences, money expenses and effort are perfect substitutes from the agent's viewpoint. But spending money for gathering knowledge is neutral from the incentives' viewpoint, whereas spending effort involves a substitution effect making the incentive constraint with respect to a harder to satisfy. Whether the principal deters or imposes to the agent to gather knowledge of type j depends on the relative magnitude of this substitution effect, as compared to the complementarity effort due to communication.

Besides, whenever the complementarity effect prevails, and the principal has not a very large competitive advantage in acquiring knowledge (Δ is not too large), it is efficient to let the agent acquire all the knowledge used in production, so as to make the largest possible the impact of knowledge on communication, and, thereby, on incentives.

5 Related Literature

Dewatripont and Tirole (2005), Liberti and Petersen (2018) and Stein (2002) all take the view that the degree of softness of information and knowledge is endogenous, and depends on costly communication activities made by the parties in a contractual relationship. For simplicity, in this paper we focus on costly communication activities undertaken by the agent only. Our analysis hinges on the assumption that in order to make transferable (harden) his soft knowledge, the agent needs to perform costly activities that involve the production and/or the use of hard knowledge.

The seminal article by Holmstrom and Milgrom (1991) provides the technology to investigate incentive issues within a multiple tasking environment. Within this framework a large literature emphasizes the "substitution effect" between costly efforts exerted in different activities. This effect arises because those "efforts" exerted to perform different activities are seen as homogeneous, depletable inputs. In our paper, we develop a variant of the multi-task model by Holmstrom and Milgrom and introduce within this set-up information gathering and communication in order to focus on the case of professional and entrepreneurial activities. Within our set-up, the main inputs supplied by the agent in production are not homogenous effort, but soft and hard knowledge and information. The non depletability of these inputs generates a strategic complementarity in the communication function, which turns out to be the main ingredient of our analysis.

The second best literature on optimal non monetary payments and the public economics literature on incentives (Atkinson and Stiglitz 1976, Bennardo et al. 2013, Fabbri and Menichini 2010, 2016, Greenwald and Stiglitz 1986, Marino and Zábojník 2008, Rajan and Wulf 2006) has shown that it may be optimal from a second best perspective to induce the agent to over-consume (under-consume) goods or production factors which reduce the marginal disutility of effort or increase the marginal productivity of effort. In this paper we show that the complementarity between soft and hard knowledge in communication, which is due to the non-depletability of the latter, implies that the second best contract imposes the agent to overproduce hard knowledge.

Holmstrom and Milgrom (1991), Rajan and Zingales (1998), and Holmstrom (1999) argue that a firm can regulate (limit) access to some key assets in order to provide better incentives to agents. Within our set-up, the principal "regulates" access to hard knowledge by imposing the agent to acquire more knowledge than he would freely do: extra access is imposed, instead of being forbidden. Our paper is also related with the organization literature focusing on delegation and integration (centralization) decisions within firms (Aghion and Tirole 1997, Dessein 2002, Dessein et al. 2010, Friebel and Raith 2010 among others).[4] In most of those papers, however, monetary incentives are muted by assumption, information resides on the agent (e.g. cannot be hardened) and "exogenous" technological motives (for instance, gains from specialization) imply that full delegation may not be efficient in the second best. In our paper, contracts are complete and monetary incentives are therefore optimally used, soft information can be hardened, albeit at a cost, and finally, whether delegation is constrained efficient only depends on an incentive trade-off.

From an empirical perspective, the literature on the role of venture capital (Da Rin et al. 2013, Hellman and Puri 2002a, 2002b, Kaplan and Stromberg 2001) shows that venture capitalists provide professionalization and business competence to the firms they fund. Consistently with our results, venture capitalists get decision power within the firms they fund and use this decision power to "impose" the acquisition of those competencies

[4]There also exists a literature starting with the seminal contributions by Marshack and Radner (1972) on optimal teams that studies optimal communication mechanisms within teams. This literature, however abstracts from incentive issues.

(knowledge). This is broadly in line with the results of our model, once the venture capital and the firm, are seen, as suggested by Kaplan and Stromberg (2001), as the parties of Principal-Agent relationship.

Finally, there exists a relatively large literature investigating agents' incentives to acquire information before contracting (Bergmann and Välimäki 2003, Crémer et al. 1998, Dasgupta and Stiglitz 1980, Hirshleifer 1971). A result common to most of that literature is that agents acquire "too much" information with respect to the second best, due to a negative information externality. In this paper, we obtain the opposite result: in the presence of full delegation, an agent acquire less information than the second best level.

6 Concluding remarks

A recent strands of papers in the literature on contracts and institutions has investigated whether incentives to gather and communicate information can shape organizational and contractual forms that are commonly observed in the real world.

In this paper, we contribute to this literature by studying a multiple tasking model where a communication technology is introduced that harden and transmit soft information from the agent to the principal. Within this set-up we show that the second best contract entails over acquisition of hard knowledge by the agent, and is precluded by full delegation.

A few simplifying assumptions are introduced to obtain this result in a stark form. In particular, we postulate that acquisition of hard and soft knowledge are the only productive activities undertaken by the agent, and that hard knowledge is the only input of the communication technology. We also assume that only hard information can be communicated to the principal.

We plan to develop the analysis of the paper by considering bilateral communication between the principal and several agents within organizations.

We see the results of this paper as a very initial step toward a better understanding of some important issues concerning the process of hardening soft knowledge within organizations and its integration with the analysis of incentives performed within standard principal agent models.

References

P. Aghion and J. Tirole, “Formal and Real Authority in Organizations”, *The Journal of Political Economy*, **105**(1), 1–29 (1997).

K. J. Arrow, “The Limits of Organization”, W. W. Norton & Company, New York (1974).

A. B. Atkinson and J. Stiglitz, “The design of tax structure:direct versus indirect taxation”, *Journal of Public Economics*, **6**(2), 55–75 (1976).

A. Bennardo, P.-A. Chiappori and J. Song, “Perks as Second Best Optimal Compensations”, *CSEF Working Papers*, No. 244, 1–49 (2013).

D. Bergemann and J.Välimäki, “Dynamic Common Agency”, *Journal of Economic Theory*, **111**(2), 23–48 (2003).

V. P. Crawford and J. Sobel, “Strategic Information Transmission”, *Econometrica*, **50**(6), 1431–1451 (1982).

P. Bolton and M. Dewatripont, “Contract Theory”, MIT Press: Cambridge and London (2005).

J. Crémer, F. Khalil and J.-C. Rochet, “Contracts and Productive Information Gathering”, *Games and Economic Behavior*, **25**(2), 174–193 (1998).

M. Da Rin, T. Hellman and M. Puri, “A Survey of Venture Capital Research”, in: George M. Constantinides, Milton Harris, Rene M. Stulz Culyer A and Newhouse J (eds.), *Handbook of the Economics of Finance* Vol. 2, Part A, 573 – 648 (2013).

P. Dasgupta and J. Stiglitz, “Industrial Structure and the Nature of Innovative Activity”, *The Economic Journal*, **11**, 266–293 (1980).

W. Dessein, “Authority and Communication in Organizations”, *The Review of Economic Studies*, **69**(4), 811–838 (2002).

W. Dessein, L. Garicano and R. Gertner, “Organizing for Synergies”, *American Economic Journal: Microeconomics*, **2**(4), 77–114 (2010).

M. Dewatripont and J. Tirole, “Modes of Communication”, *Journal of Political Economy*, **113**(6), 1217–1238 (2005).

D. Fabbri and A. M. C. Menichini, “Trade credit, collateral liquidation, and borrowing constraints”, *Journal of Financial Economics*, **96**(3), 413–432 (2010).

D. Fabbri and A. M. C. Menichini, “The Commitment Problem of Secured Lending”, *Journal of Financial Economics*, **120**(3), 561–584 (2016).

G. Friebel and M. Raith, "Resource Allocation and Organizational Form", *American Economic Journal: Microeconomics*, **2**(2), 1–33 (2010).

L. Garicano and Y. Wu, "Knowledge, Communication, and Organizational Capabilities", *Articles in Advance*, 1–16 (2012).

B. C. Greenwald and J. Stiglitz, "Externalities in Economies with Imperfect Information and Incomplete Markets", *The Quarterly Journal of Economics*, **101**(2), 229–264 (1986).

F. A. Hayek, "The use of knowledge in society", *American Economic Review*, **35**(4), 519–530 (1945).

T. Hellmann and M. Puri, "Venture Capital and the Professionalization of Start-Up Firms: Empirical Evidence", *The Journal of Finance*, **57**(1), 169–197 (2002a).

T. Hellmann and M. Puri, "On the Fundamental Role of Venture Capital", *Economic Review*, **87**(4), 19–23 (2002b).

J. Hirshleifer, "The Private and Social Value of Information and the Reward to Inventive Activity", *American Economic Review*, **61**, 561–574 (1971).

B. Holmstrom, "Moral hazard and observability", *The Bell Journal of Economics*, **10**(1), 74–91 (1979).

B. Holmstrom and P. Milgrom, "Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design", *Journal of Law, Economics, & Organization*, **7**, 24–52 (1991).

B. Holmstrom, "The firm as a subeconomy", *Journal of Law Economics & Organization*, **15**(1), 74–102 (1999).

M. C. Jensen and W. H. Heckling, "Specific and General Knowledge, and Organizational Structure", *Journal of Applied Corporate Finance*, **8**(2), 4–18 (1995).

S. N. Kaplan and P. Stromberg, "Venture Capitals As Principals: Contracting, Screening, and Monitoring", *The American Economic Review*, **91**(2), 426–430 (2001).

J. M. Liberti, M. A. Petersen, "Information: Hard and Soft", *NBER Working Paper*, No. **25075** (2018).

A. M. Marino and J. Zábojník, "Work-related perks, agency problems, and optimal incentive contracts", *The RAND Journal of Economics*, **39**(2), 565–585 (2008).

J. Marschak and R. Radner, “Economic Theory of Teams”, Cowles Foundation for Research in Economics at Yale University, Monograph 22, Yale University Press, New Haven and London (1972).

P. Milgrom, “Good News and Bad News: Theorems and Applications”, *The Bell Journal of Economics*, **12**(2), 380–391 (1981).

M. Polanyi, “The Tacit Dimension”, Anchor Books, GardenCity, NY (1966).

M. Raith “Specific knowledge and performance measurement”, *The RAND Journal of Economics*, **39**(4), 1059–1079 (2008).

R. Rajan and L. Zingales, “Power in a Theory of the Firm”, *The Quarterly Journal of Economics*, **113**(2), 387–432 (1998).

R. Rajan and J. Wulf, “Are perks purely managerial excess”, *Journal of Financial Economics*, **79**(1), 1–33 (2006).

J.C. Stein, “Information Production and Capital Allocation: Decentralized vs. Hierarchical Firms”, *Journal of Finance*, **57**(5), 1891–1921 (2002).

Appendix

Proof of Proposition 3.1.

(I) Suppose $(s_\theta^{SB}, s_{\omega_2}^{SB}) \neq 0$. Consider the contract

$$\gamma' := \{a^{SB}, z_A^{SB}, z_P^{SB}, t', s_c^{SB}, s_{\omega_1}^{SB}, s'_\theta, s'_{\omega_2}\} \tag{12}$$

where $(s'_\theta, s'_{\omega_1}) = 0$ and $t' = t^{SB} + (s_{\omega_2}^{SB} + s_\theta^{SB})\mu_{\omega_2} - \varepsilon$, with $\varepsilon > 0$ and small enough. Notice that, under the new contract, a^{SB} is still incentive compatible since

$$\frac{\partial E[U(\gamma)]}{\partial a} = (s_c + s_{\omega_1})\frac{\partial h(a)}{\partial a} - \frac{\partial \varphi(a)}{\partial a} - \frac{1}{2}\eta\left(s_c^2 \frac{\partial \sigma_c^2(a, z_A)}{\partial a}\right) \tag{13}$$

is independent from s_θ and s_{ω_2}.

Moreover, at γ', both the principal and the agent are better off. The principal's expected profit increases by ε since a^{SB} is incentive compatible at the new contract, and the agent's expected utility increases by

$$\frac{1}{2}\eta(s_\theta^2\sigma_\theta^2(z) + s_{\omega_2}^2\sigma_{\omega_2}^2 + 2s_\theta s_{\omega_2}\sigma_{\theta\omega_2}(z)) - \varepsilon\Big|_{\substack{s_\theta = s_\theta^{SB} \\ s_{\omega_2} = s_{\omega_2}^{SB}}} \tag{14}$$

This contradicts the optimality of $(s_\theta^{SB}, s_{\omega_2}^{SB}) \neq 0$.

(II) First, incentive compatibility implies that for the agent to choose $a > 0$, at least one component of the vector $s = (s_c, s_{\omega_1})$ must be strictly positive.

We start by showing that $s_c \geq 0$ and $s_{\omega_1} \geq 0$, with at least one strict inequality.

Suppose by the contrary that $s_c^{SB} > 0$ and $s_{\omega_1}^{SB} < 0$. Take the contract $\gamma' := \{a', z_A^{SB}, z_P^{SB}, t', s_c^{SB}, s'_{\omega_1}\}$ with $s'_{\omega_1} = 0$ and $t' = t^{SB} + s_{\omega_1}^{SB}(h(a') + \mu_{\omega_1}) - \varepsilon$, with $\varepsilon > 0$ and small enough, and

$$a'(z_A, s_c) = \arg\max_a \quad E[U(\gamma)]\big|_{s_{\omega_1}=0} \tag{15}$$

Notice that

$$\frac{\partial E[U(\gamma)]}{\partial a}\Big|_{s_{\omega_1}=0} = s_c\frac{\partial h(a)}{\partial a} - \frac{\partial \varphi(a)}{\partial a} - \frac{1}{2}\eta s_c^2\frac{\partial \sigma_c^2(a, z_A)}{\partial a} \tag{16}$$

implies $a' > a^{SB}$. Hence, at the contract γ', the principal is better off than at γ^{SB} because the agent's expected payment is reduced by ε and his effort is increased. The agent is also better off for the two following reasons.

First, a revealed preferences argument implies that at γ' the agent cannot be better off by choosing a^{SB} instead of a'. Second,

$$\frac{1}{2}\eta(s_{\omega_1}^2\sigma_{\omega_1}^2 + 2s_c s_{\omega_1}\sigma_{c\omega_1}(a^{SB}, z_A^{SB})) - \varepsilon\Big|_{\substack{s_c = s_c^{SB} \\ s_{\omega_1} = s_{\omega_1}^{SB}}} > 0 \tag{17}$$

for all $\rho_{c\omega_1} \leq 0$. And, by continuity, there exists a positive value $\bar{\rho}$ of $\rho_{c\omega_1}$ such that (17) holds true for all $\rho_{c\omega_1} \leq \bar{\rho}$.

Moreover, notice that, by the same logic as above, one could not have $s_c^{SB} < 0$ and $s_{\omega_1}^{SB} > 0$ at the second best.

Finally, since x_c is not a sufficient statistics of ω_1, nor ω_1 is a sufficient statistics of x_c, one must have $s_c > 0$ and $s_{\omega_1} > 0$.

(III) The first-order conditions w.r.t. z_A and z_P are

$$\frac{\partial E[R(\gamma)]}{\partial z_A} + \lambda_1 \left(-\frac{1}{2}\eta s_c^2 \frac{\partial^2 \sigma_c^2(a, z_A)}{\partial a \partial z_A} \right) + \\ + \lambda_2 \left(-q_A - \frac{1}{2}\eta \left(s_c^2 \frac{\partial \sigma_c^2(a, z_A)}{\partial z_A} + s_\theta^2 \frac{\partial \sigma_\theta^2(z)}{\partial z_A} + 2 s_\theta s_{\omega_2} \frac{\partial \sigma_{\theta\omega_2}(z)}{\partial z_A} \right) \right) \leq 0 \tag{18}$$

and,

$$\frac{\partial E[R(\gamma)]}{\partial z_P} - q_P + \lambda_2 \left(-\frac{1}{2}\eta \left(s_\theta^2 \frac{\partial \sigma_\theta^2(z)}{\partial z_P} + 2 s_\theta s_{\omega_2} \frac{\partial \sigma_{\theta\omega_2}(z)}{\partial z_P} \right) \right) \leq 0 \tag{19}$$

Since the marginal utility of income is the same for the agent and the principal in the second best (because income enters linearly in the certainty equivalent of the agent), the lagrange multiplier λ_2 is equal to 1 at the second best. It is then immediate to verify that for the inequality in (18) to hold, the inequality in (19) must hold strictly whenever $q_A \leq q_P + \Delta$, with Δ small enough.

Proof of Proposition 3.2.

Consider the first-order condition of the second best program w.r.t. z_A

$$\frac{dE[\Pi(\gamma^{SB})]}{dz_A} + \lambda \frac{\partial}{\partial z_A} \left(\frac{dE[U(\gamma^{SB})]}{da} \right) + \mu \frac{dE[U(\gamma^{SB})]}{dz_A} = 0 \tag{20}$$

Now, note that

$$\frac{dE[U(\gamma^{SB})]}{dz_A} = \frac{\partial E[U(\gamma^{SB})]}{\partial a} \frac{\partial a}{z_A} + \frac{\partial E[U(\gamma^{SB})]}{\partial z_A} \tag{21}$$

Moreover, incentive compatibility implies $\frac{\partial E[U(\gamma^{SB})]}{\partial a} = 0$. Hence,

$$\frac{dE[U(\gamma^{SB})]}{dz_A} = \frac{\partial E[U(\gamma^{SB})]}{\partial z_A} \tag{22}$$

In addition,

$$\frac{\partial}{\partial z_A} \left(\frac{dE[U(\gamma^{SB})]}{da} \right) = -\frac{1}{2}\eta s_c^2 \frac{\partial^2 \sigma_c^2(a, z_A)}{\partial a \partial z_A} > 0 \tag{23}$$

under our complementarity assumption $\left(\frac{\partial^2\sigma_c^2(a,z_A)}{\partial a\partial z_A} < 0\right)$.

In the following, we will show $\frac{dE[\Pi(\gamma^{SB})]}{dz_A} \geq 0$ which, together with (22) and (23), implies $\frac{dE[U(\gamma^{SB})]}{dz_A} < 0$, and hence $z_A^{SB} > z_A^*$.

To this end, we begin by observing that $\frac{dE[\Pi(\gamma^{SB})]}{dz_P} \leq 0$. This simply comes from the first-order condition of the second-best program (7) w.r.t. z_P:

$$\frac{dE[\Pi(\gamma^{SB})]}{dz_P} + \lambda_1 \frac{\partial}{\partial z_P}\left(\frac{dE[U(\gamma^{SB})]}{da}\right) + \lambda_2 \frac{dE[U(\gamma^{SB})]}{dz_P} = 0 \tag{24}$$

since z_P does not affect directly neither $\frac{dE[U(\gamma^{SB})]}{da} = 0$, nor $\frac{dE[U(\gamma^{SB})]}{dz_P}$.

Finally, we show that $\frac{\partial E[\Pi(\gamma^{SB})]}{\partial z_A} \geq 0$. The proof is by contradiction. Assume $\frac{\partial E[\Pi(\gamma^{SB})]}{\partial z_A} < 0$. Notice that $\frac{\partial E[\Pi(\gamma^{SB})]}{\partial z_P} < \frac{\partial E[\Pi(\gamma^{SB})]}{\partial z_A}$ because the principal does not pay for the cost of z_A. Then, one must necessarily have $\frac{\partial E[\Pi(\gamma^{SB})]}{\partial z_P} < 0$. But this implies that z_P^{SB} must necessarily equal to zero by second-best optimality. In turn, if $z_P^{SB} = 0$, then $\frac{\partial E[\Pi(\gamma^{SB})]}{\partial z_A} \geq 0$.

Proof of Proposition 3.3.

Let $\gamma' = (a^{FD}, z_A^{FD} + dz_A, z_P^{FD}, t^{FD} + dt, s')$ where dz_A is positive and small enough, and

$$s' = (s_c^{FD}, s_{\omega_1}^{FD} + ds_{\omega_1}, s_\theta^{FD}, s_{\omega_2}^{FD}),$$

$$dt = -\left(\left(\frac{\partial E[U(\gamma^{FD})]}{\partial s_{\omega_1}}\right)\left(\frac{1}{2}\eta s_c^2 \frac{\partial^2\sigma_c^2(a,z_A)}{\partial a\partial z_A}\left(\frac{\partial h(a)}{\partial a}\right)^{-1}\right) + \left(\frac{\partial E[U(\gamma^{FD})]}{\partial z_A}\right)\right) dz_A,$$

$$ds_{\omega_1} = \frac{1}{2}\eta s_c^2 \frac{\partial^2\sigma_c^2(a,z_A)}{\partial a\partial z_A}\left(\frac{\partial h(a)}{\partial a}\right)^{-1} dz_A. \tag{25}$$

Let

$$dU \approx E[U(\gamma')] - E[U(\gamma^{FD})] \tag{26}$$

and

$$dIC_a \approx \frac{\partial E[U(\gamma')]}{\partial a} - \frac{\partial E[U(\gamma^{FD})]}{\partial a} \tag{27}$$

be the first order approximations of the difference between the expected utility evaluated at γ' and γ^{FD}, and of the difference between the value of the incentive constraint evaluated at γ' and γ^{FD} respectively.

As an intermediary step, we show that $dIC_a \approx 0$, $dU \approx 0$ at $\gamma = \gamma'$, and $E[\Pi(\gamma')] > E[\Pi(\gamma^{FD})]$.

By totally differentiating $\frac{\partial E[U(\gamma')]}{\partial a}$ w.r.t. (ds_{ω_1}, dz_A) one obtains

$$dIC_a = \frac{\partial h(a)}{\partial a} ds_{\omega_1} - \frac{1}{2}\eta s_c^2 \frac{\sigma_c^2(a, z_A)}{\partial a \partial z_A} dz_A = 0 \tag{28}$$

Now, denote

$$\chi(s, \sigma_c^2, \sigma_{\omega_1}^2 \sigma_{c\omega_1}) = -\frac{1}{2}\eta \left(s_c^2 \sigma_c^2(a, z_A) + s_{\omega_1}^2 \sigma_{\omega_1}^2 + 2 s_c s_{\omega_1} \sigma_{c\omega_1}(a, z_A)\right) \tag{29}$$

By totally differentiating $EU(\gamma^{FD})$ w.r.t. (ds_{ω_1}, dz_A) one obtains

$$dU = (h(a) + \mu_{\omega_1})\, ds_{\omega_1} - q_A dz_A + \frac{\partial \chi(\cdot)}{\partial s_{\omega_1}} ds_{\omega_1} + \frac{\partial \chi(\cdot)}{\partial z_A} dz_A = 0 \tag{30}$$

Finally, totally differentiating $E[\Pi(\gamma^{FD})]$ w.r.t. dz_A, and then replacing the values of dt and ds_{ω_1} in (25), one obtains

$$\left(\frac{1}{2}\eta s_c^2 \frac{\partial^2 \sigma_c^2(a, z_A)}{\partial a \partial z_A} \left(\frac{\partial h(a)}{\partial a}\right)^{-1} \left(\frac{\partial \chi(\cdot)}{\partial s_{\omega_1}}\right) + \right.$$
$$\left. + \frac{\partial E[\Pi(\gamma^{FD})]}{\partial z_A} + \frac{\partial E[U(\gamma^{FD})]}{\partial z_A} \right) dz_A \tag{31}$$

and it is straightforward to verify that this expression is strictly positive for $dz_A > 0$.

A continuity argument then implies that there is a contract γ'' that satisfies the participation and incentive constraint w.r.t. a, and yields larger expected profit to the principal than γ^{FD}.

Now consider the following program

$$\max_{\gamma} E[\Pi(\gamma)] \quad s.t. \quad IC_a(\gamma) = 0;\ E[U(\gamma) = \bar{U}];\ z_A = \bar{z}_A \tag{32}$$

Let $\hat{\Pi}(\bar{z}_A)$ be the value function associated to this program. The previous result implies that $\frac{\partial \hat{\Pi}(\bar{z}_A)}{\partial \bar{z}_A}\Big|_{\bar{z}_A = z_A^{FD}} > 0$. The reason is the following. First, $\hat{\Pi}(\bar{z}_A)\Big|_{\bar{z}_A = z_A^{FD}} = E[\Pi(\gamma^{FD})]$ by construction. Moreover, there is a feasible contract γ'' imposing $z_A^{FD} + dz_A$ with $dz_A > 0$ yielding a larger profit than γ^{FD}. Therefore the value function associated to program (32) must necessarily take a larger value in $z_A^{FD} + dz_A$ than in z_A^{FD}. Hence, if the program in (32) has an unique local maximum than $z^{FD} < z^{SB}$. If, instead, the second-best program has several local maxima, there is at least one of them such that $z^{FD} < z^{SB}$.

Proof of Proposition 4.1.

We first define the agent's certainty equivalent of the expected utility and the principal's expected profit in the extended scenario with multiple knowledge types; subsequently, the three statements in Proposition 4.1 are proved.

The certainty equivalent is

$$\mathbf{CE}(\gamma) = t+h(a)(s_c+s_{\omega_1})+s_{\omega_1}\mu_{\omega_1}+(s_\theta+s_{\omega_2})\mu_{\omega_2}-\varphi\left(a+\sum_{j=1}^{J}e_j(z_{jA})\right)+$$

$$-\sum_{j=1}^{J}q_{jA}z_{jA}-\frac{1}{2}\eta\left(s_c^2\sigma_c^2(a,\mathbf{z_A})+s_{\omega_1}^2\sigma_{\omega_1}^2+2s_cs_{\omega_1}\sigma_{c\omega_1}(a,\mathbf{z_A})+\right.$$

$$\left.+s_\theta^2\sigma_\theta^2(\mathbf{z})+s_{\omega_2}^2\sigma_{\omega_2}^2+2s_\theta s_{\omega_2}\sigma_{\theta\omega_2}(\mathbf{z})\right) \quad (33)$$

where $\mathbf{z}=(\mathbf{z_A},\mathbf{z_P})$.

The expected profit of the principal in the extended set-up is

$$\mathbf{E}[\mathbf{\Pi}(\gamma)] = \int_{-\infty}^{+\infty}\int_{0}^{+\infty} g_1(x+\omega_1)f(\omega_1|x_c,a,\mathbf{z_A})f(x_c|a,\mathbf{z_A})d\omega_1dx_c+$$

$$+\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\left(g_2(m^*(\theta,\mathbf{z})+\omega_2)-m^*(\theta,\mathbf{z})\right)f(\omega_2|\theta,\mathbf{z})f(\theta|\mathbf{z})d\omega_2d\theta+$$

$$-\left(t+h(a)(s_c+s_{\omega_1})+s_{\omega_1}\mu_{\omega_1}+(s_\theta+s_{\omega_2})\mu_{\omega_2}\right)-\sum_{j=1}^{J}q_{jP}z_{jP} \quad (34)$$

(I) In the presence of multiple knowledge types, the first-order condition w.r.t. z_{jA} is

$$\frac{d\mathbf{E}[\mathbf{\Pi}(\gamma^{\mathbf{SB}})]}{dz_{jA}}+\lambda_1\frac{\partial}{\partial z_{jA}}\left(\frac{d\mathbf{E}[\mathbf{U}(\gamma^{\mathbf{SB}})]}{da}\right)+\lambda_2\frac{d\mathbf{E}[\mathbf{U}(\gamma^{\mathbf{SB}})]}{dz_{jA}}=0 \quad (35)$$

where

$$\frac{\partial}{\partial z_{jA}}\left(\frac{d\mathbf{E}[\mathbf{U}(\gamma^{\mathbf{SB}})]}{da}\right)=-\frac{1}{2}\eta s_c^2\frac{\partial^2\sigma_c^2(a,\mathbf{z_A})}{\partial a\partial z_{jA}}-\frac{\partial^2\varphi(\cdot)}{\partial a\partial e_j}\frac{\partial e_j(z_{jA})}{\partial z_{jA}} \quad (36)$$

If $j\in J'$, then $\frac{\partial e_j(z_{jA})}{\partial z_{jA}}=0$ and, in turn, $\frac{\partial}{\partial z_{jA}}\left(\frac{d\mathbf{E}[\mathbf{U}(\gamma^{\mathbf{SB}})]}{da}\right)>0$ since the sole complementarity effect survives in (36).

The rest of the proof follows the same logic as the proof of Proposition 3.2.

(II) From Proposition 3.1, if $q_{jA}\leq q_{jP}+\Delta$, with $\Delta>0$, then $z_{jP}^{SB}=0$ and, in turn, $\frac{\partial\mathbf{E}[\mathbf{\Pi}(\gamma^{\mathbf{SB}})]}{\partial z_{jA}}\geq 0$. In addition, since $\frac{\partial}{\partial z_{jA}}\left(\frac{d\mathbf{E}[\mathbf{U}(\gamma^{\mathbf{SB}})]}{da}\right)$ is assumed to be positive, then $\frac{d\mathbf{E}[\mathbf{U}(\gamma^{\mathbf{SB}})]}{dz_{jA}}<0$, which implies $z_{jA}^*<z_{jA}^{SB}$ for all

$j \in J''$ such that $\frac{\partial}{\partial z_{jA}} \left(\frac{d\mathbf{E}[\mathbf{U}(\gamma^{\mathbf{SB}})]}{da} \right) > 0$.

(III) First, note that $(s_\theta^{SB}, s_{\omega_2}^{SB}) = 0$, since the logic of the proof of the first statement in Proposition 3.1 applies equivalently to the extended model.

Now, consider the first-order conditions w.r.t. z_{jA} and z_{jP}, respectively,

$$\frac{\partial \mathbf{E}[\mathbf{R}(\gamma)]}{\partial z_{jA}} + \lambda_1 \left(-\frac{1}{2} \eta s_c^2 \frac{\partial^2 \sigma_c^2(a, \mathbf{z_A})}{\partial a \partial z_{jA}} \right) + \\ + \lambda_2 \left(-q_{jA} - \frac{1}{2} \eta s_c^2 \frac{\partial \sigma_c^2(a, \mathbf{z_A})}{\partial z_{jA}} \right) \leq 0 \quad (37)$$

and,

$$\frac{\partial \mathbf{E}[\mathbf{R}(\gamma)]}{\partial z_{jP}} - q_{jP} \leq 0 \quad (38)$$

where $\mathbf{E}[\mathbf{R}(\gamma)]$ is the expected revenue function in the presence of multiple knowledge types.

Let $\Delta > 0$ be an increase of q_{jA} such that the left-hand side of (37) is equal to the left-hand side of (38). Since $\lambda_2 = 1$ from the proof of Proposition 3.2, if $q_{jP} < q_{jA} + \Delta$, then the inequality in (37) must hold strictly, implying $z_{jA}^{SB} = 0$ and, in turn, $z_{jP}^{SB} > 0$. This proves the first statement in (III).

To prove the second part of (III), consider, once again, the inequalities in (37) and (38), and let $\Delta > 0$ be defined as above. If $q_{jP} > q_{jA} + \Delta$, then the inequality in (38) must hold strictly, implying $z_{jP}^{SB} = 0$ and, in turn, $\frac{\partial E[\Pi(\gamma)]}{\partial z_{jA}} \geq 0$. Hence, in the presence of multiple knowledge types, $\frac{\partial E[\Pi(\gamma)]}{\partial z_{jA}}$ and $\frac{\partial}{\partial z_{jA}} \left(\frac{d\mathbf{E}[\mathbf{U}(\gamma^{\mathbf{SB}})]}{da} \right)$ are not uniquely signed in equation (35). In addition, if q_{jA} is small enough, then $\frac{d\mathbf{E}[\mathbf{U}(\gamma^{\mathbf{SB}})]}{dz_{jA}} > 0$, meaning that the substitution effect ($\frac{\partial}{\partial z_{jA}} \left(\frac{d\mathbf{E}[\mathbf{U}(\gamma^{\mathbf{SB}})]}{da} \right) < 0$) is dominating, and, in turn, $z_{jA}^* > z_{jA}^{SB}$.