UNIVERSITÀ CATTOLICA DEL SACRO CUORE

DIPARTIMENTO DI ECONOMIA INTERNAZIONALE
DELLE ISTITUZIONI E DELLO SVILUPPO



Ferdinando Colombo and Guido Merzoni


# Stable delegation in an unstable environment


N. 0701



V&P

UNIVERSITÀ CATTOLICA DEL SACRO CUORE

DIPARTIMENTO DI ECONOMIA INTERNAZIONALE
DELLE ISTITUZIONI E DELLO SVILUPPO

Ferdinando Colombo and Guido Merzoni

# Stable delegation in an unstable environment

N. 0701

V&P

I Quaderni del Dipartimento di Economia internazionale delle istituzioni e dello sviluppo possono essere richiesti alla Segreteria: (Tel. 02/7234.3788 - Fax 02/7234.3789 - E-mail: segreteria.diseis@unicatt.it).
www.unicatt.it/dipartimenti/diseis

Università Cattolica del Sacro Cuore, Via Necchi 5 - 20123 Milano

www.vitaepensiero.it


ISBN 978-88-343-1647-4

# Stable delegation in an unstable environment *

Ferdinando Colombo † Guido Merzoni ‡

November 15, 2007

**Abstract**

The Kreps–Wilson–Milgrom–Roberts framework is one of the most renowned ways of modelling reputation–building. Once the number of repetitions of the game is considered as a choice variable, such a framework can fruitfully be employed to study the optimal length of a relationship. We analyze a model where a principal delegates to an agent the task of playing with a third party a finitely repeated trust game, characterize the optimal length of the relationship between principal and agent when the principal's preferences on the agent's type stochastically change over time and show that stable relationships may optimally obtain (even) in very unstable environments.

*Keywords:* repeated games with incomplete information, reputation, stable relationships, changing environment.

*J.E.L. Classification:* C72, D82, L14.

---

* We wish to thank Carlo Beretta, Luigi Curini, Carla Peri, Giulio Piccirilli, Jonathan Thomas, T. Erika Uberti and Salvatore Vassallo, as well as two anonymous referees, for useful suggestions and comments. We have also benefited from comments on a related paper by seminar participants at Università Cattolica del Sacro Cuore, Milan, E.A.R.I.E., Helsinki and the S.E.T. conference on "Organization Design, Corporate Governance and Regulation", Milan. The usual disclaimer applies. Ferdinando Colombo acknowledges the financial support of MIUR under the Research Project "Agenti eterogenei, fragilità finanziaria e dinamiche complesse". Guido Merzoni acknowledges the financial support of MIUR under the Research Project "Dinamica strutturale: imprese, organizzazioni, istituzioni".

† *Corresponding author.* ITEMQ, Faculty of Economics, Università Cattolica del Sacro Cuore, L.go Gemelli, 1, 20123 Milano. E–mail: `ferdinando.colombo@unicatt.it`.

‡ DISEIS, Faculty of Political Sciences, Università Cattolica del Sacro Cuore, L.go Gemelli, 1, 20123 Milano. E–mail: `guido.merzoni@unicatt.it`.

Ferdinando Colombo and Guido Merzoni

**Stable delegation in an unstable environment**

n. 0701

# 1 Introduction

Delegation is ubiquitous in most human activities. Many important issues related to delegation, such as the allocation of decisional power (Aghion and Tirole, 1997), the existence of conflicts of interests (Holmstrom, 1979; Grossman and Hart, 1983; Mirrlees, 1999) and the strategic design of incentive contracts (Fershtman and Judd, 1987) have been widely explored in the literature. There is, however, another important issue that has so far received relatively little attention, namely the time length of delegation arrangements which, in our opinion, should depend on both the characteristics of the delegated activity and the environment where such an activity is carried out. Even a cursory look at the term–length of central bankers in different countries or the duration of labor contracts across different industries seems to confirm our opinion.

In this paper, we focus on an important trade–off that characterizes the choice of the optimal length of delegation arrangements when, as it is often the case, the delegated activity involves moral hazard and is carried out in a changing environment: on the one hand, reputation building, which may be helpful in solving the moral hazard problems, requires sufficiently stable relationships, while, on the other hand, the need to promptly react to a changing environment is favored by flexible delegation arrangements.

We model the first element of the above trade–off by assuming that a principal delegates to an agent the task of playing with a third party a repeated trust game, *i.e.*, a stylized representation of an interaction characterized by moral hazard. [1] In similar settings, the effectiveness of long–term relationships as a way of increasing efficiency when third party's enforcement is unavailable has been recognized for long [2], mainly modelling reputation building in an infinitely repeated game. However, in such models, the important link between the *degree* of stability of a relationship and the *amount* of cooperation within such a relationship [3] cannot be easily characterized.

[1] The trust game was first presented as such by Kreps (1990). In recent years, it has also been the object of many experimental investigations, such as Anderhub et al. (2002), Berg et al. (1995) and Burnham et al. (2000).

[2] For example, Barro and Gordon (1983), Backus and Driffill (1985) and Barro (1986) show that ongoing interaction between a policy–maker and the private sector can mitigate the inflation bias and restore some credibility to a low–inflation monetary policy, and Bull (1987) and MacLeod and Malcomson (1989) show that in long–lasting relationships employers can induce employees to provide higher effort through the implicit promise to pay them bonuses.

[3] For example, one could reasonably expect that *ceteris paribus* the behavior of one player is the more cooperative the larger the number of times the same game will be

Moreover, the length of the relationship is taken as given, while we think that treating duration as a choice variable is appropriate and interesting, since in most conceivable situations the parties may optimally take actions that somehow affect the stability of a relationship and so also affect their ability to cooperate. [4]

For our purposes, it seems more rewarding to model the benefits of reputation-building by introducing incomplete information *à la* Kreps-Wilson -Milgrom-Roberts (Kreps and Wilson, 1982, Milgrom and Roberts, 1982) in a finitely repeated trust game, and to assume that the principal, by choosing the length of the delegation arrangement with her agent, also selects the number of repetitions of the trust game between her agent and a third party. [5] This modelling strategy allows us to meaningfully treat the duration of a relationship as a choice variable and to reach quite neat and "reasonable" conclusions on the link between the stability of a relationship and the amount of cooperation within such a relationship: trust cannot arise unless the parties perceive their relationship to be sufficiently stable, a more stable relationship gives rise to more trust and the larger the number of periods where one party honored trust the better his reputation for trustworthiness.

The simplifying assumption that the principal is able to choose the actual length of the delegation arrangement with her agent allows us to capture the much more general idea that the principal is able to affect the degree of stability of the relationship between the delegate and other parties with whom he interacts on her behalf. This assumption implies that the delegation arrangement is equipped with the commitment value that may make long–term relationships preferable to short–term ones, because they allow at least one party to follow an ex–post sub–optimal course of actions, thus obtaining a strategic advantage, as in Malcomson and Spinnewyn (1988)

---

played in the future. For some experimental evidence supporting this intuition see, *e.g.*, Tables **8.22** and **8.23** in Camerer (2003) (and the paper to which they refer, Brandt and Figueras, 2003).

[4] Such actions may range from constitutionally setting a term length for public officials, to the choice of the duration and the degree of flexibility of employment contracts or purchasing agreements, to relationship–specific investments.

[5] To the best of our knowledge, in its relatively few applications, the Kreps–Wilson–Milgrom–Roberts framework has never been used to study optimal stability, because the number of repetitions of the game either had no particular meaning or was taken as given. For example, Backus and Driffill (1985, p. 532) and Barro (1986, p. 3) apparently use this type of models mainly because they yield a unique equilibrium. In fact, they also recognize that there may be a rationale for a finite horizon (government term, with no possibility of re–election; policy–maker term), but do not consider the number of repetitions of the game as a choice variable.

and Fudenberg *et al* (1990). In fact, when the parties have an incentive to renegotiate, stability can be signaled through the length of the arrangement only if the principal is able to build a reputation for not renegotiating such length; assuming non–renegotiability can therefore be viewed as a way to avoid modelling a second kind of reputation–building. As for the assumption that the principal cannot re–hire the same agent, in Colombo and Merzoni (2006) we have shown that, under reasonable conditions, in a repeated trust game the commitment to a long–term relationship is valuable. [6] Hence, even if one allows for renewability, a sequence of short–term delegation arrangements need not be optimal in our model. Ruling out the possibility of re–hiring is therefore just a simplifying assumption that allows us to focus on the case when commitment is valuable. Finally, while in some applications the assumption that the principal is able to choose the actual length of the delegation arrangement with her agent can be taken at its face value (for example, some central bankers or the members of many independent agencies in several countries are appointed for a fixed, non–renewable term, and can be removed from their office only in rather exceptional cases), in other applications it holds only in approximate terms (for example, even if a permanent position within a firm does not necessarily imply that a worker always stays with the same firm, it seems reasonable to assume that on average the relationship between a firm and its permanent workers lasts longer than its relationship with temporary workers).

We have up to now focused on the moral hazard nature of delegated activities. Of course, in this setting, allowing the principal to choose the number of repetitions of the game does not lead us too far in the determination of the optimal length of delegation: indeed, the principal always has the incentive to indefinitely increase such **a** number. However, it is quite obvious that the commitment to a long–term relationship may have some drawbacks. Once one takes them into account, the choice of the optimal length of a relationship becomes a real and important issue.

The second effect of the length of delegation arrangements we consider is that longer delegation means less flexibility. This makes it more costly or

[6] In Colombo and Merzoni (2006) we considered a simple two–period model where a principal (directly) plays a (possibly repeated) trust game with her agent in a stable environment, the agent's skills are private information at the beginning of the relationship and later revealed to the principal, and one–period contracts can be renewed. We showed that it may be optimal for the principal to commit to a two–period relationship with her agent (renewable one–period contracts may indeed give rise to not–sufficiently–stable relationships, which inhibit the emergence of trust). In Colombo and Merzoni (2004) the same result was obtained in a similar (two–period) model where, however, as in the present paper, the playing of the trust game is delegated.

even impossible to react to new information.

We assume that agents are of different types; the type of an agent is publicly observable and verifiable. The principal lives however in a changing environment: in each period there is a given probability that a non–verifiable shock occurs, which changes her needs. Hence, after a shock, the principal would have the incentive to replace the existing agent with a new one who exactly matches her needs at that moment; however, she can replace her agent only after the agreed upon period of delegation expires. Shorter delegation allows the principal to more often choose an agent who matches her needs, and so to have for a lower (expected) number of periods an unfit agent.

The optimal length of the delegation arrangement depends on the relative strength of the benefits of a stable *vs* a flexible relationship. On the one hand, the longer the length of the arrangement, the more stable the relationship between the agent and a third party, and so the larger the amount of trust within such a relationship. On the other hand, the shorter the length of the arrangement, the more often the principal is able to replace the old agent with a new, more suitable one. We think that such a trade–off characterizes most problems of delegation. For example, long–term delegation to a conservative central banker reduces the inflation bias, but entails the risk of having for some time a central banker who has preferences that differ from those of the public (as expressed, for example, in the outcome of a general election) [7]; long–term delegation to a manager to supervise employees and pay them bonuses favors the emergence of trust within the firm, but also makes it more difficult to adjust to shocks. Our model is meant therefore to represent an encompassing framework that characterizes the optimal length of delegation arrangements as determined by the very nature of the delegated activities, which are defined by their general features and by the environment where they are carried out.

This paper can be related to the (relatively scarce) literature on optimal contract length. In a very general setting, Malcomson and Spinnewyn (1988) and Fudenberg, Holmstrom and Milgrom (1990), besides recognizing the value of commitment in long–term relationships, also provide conditions for the equivalence between short–term and long–term contracts. Yet, they do not model any explicit benefit of short–term contracts, as we do.

Gray (1978), Dye (1985), Harris and Holmstrom (1987) and Cantor (1988) take into account an advantage of short–term contracts: as time passes, the

[7] A similar problem has been studied by Waller and Walsh (1996) and Lin (1999) who, however, fail to endogenize the benefits of long–term delegation. See also Lindner (2000).

parties' acquire information, and this enables them to profitably reshape the contract. In those papers, however, the commitment to a long–term relationship has no value *per se*, so the only reason why spot contracts are not always optimal is the need to bear exogenously given writing or re–contracting costs. Our paper endogenizes the above costs as forgone returns for the principal due to the time needed by a new agent to build his reputation and so relates them to an intrinsic and important feature of repeated relationships, namely the dependence of the amount of cooperation on the perception of stability.

Our comparative statics analysis leads to two main results.

First, as expected, as long as reputation is worth–building, the optimal length of delegation increases as reputation becomes more costly to be established and decreases as it becomes more costly for the principal to keep an agent who does not fit her needs. However, when either of these costs becomes too large, the principal chooses to renounce reputation–building and hires in each period an agent who exactly matches her needs, so there is a downward jump in the optimal length of delegation. This discontinuity depends on the indivisibility of the investment in reputation in our model, which seems compatible with many real–world examples, where trust generally emerges only after a (trial) period characterized by a sub–optimal level of cooperation; when such a period is too long, it may then become optimal to give up reputation–building altogether.

Our second main comparative statics result concerns the relation between variability of the environment and optimal length of delegation. The literature on optimal contract length has studied the effects of variability and proposed two main views: according to Gray (1978), length should decrease with the size of shocks because larger shocks make the need to adjust contracts' provisions more frequent; on the other hand, Danziger (1988) claims that, as variability increases, contracts should become longer to provide insurance more efficiently. [8] In our setting, agents are risk–neutral, so the insurance argument plays no role. Thus, one could conjecture that we should obtain a decreasing relationship. We show that this is actually the case if one measures the variability of the environment by the maximum loss the principal suffers in one period if her agent does not fit her needs. However, when the variability of the environment is seen as a lack of stability and measured by the probability that a shock occurs, the relationship between variability and optimal length of delegation turns out to be decreasing only if the cost of building reputation is small relative to the loss for the principal

[8] Many empirical studies have tried to test these two conflicting predictions (*e.g.*, Christofides and Wilton, 1983; Murphy, 2000; Wallace, 2001; Rich and Tracy, 2004). Yet, so far the evidence seems mixed.

of keeping an unfit agent, whereas it is U–shaped otherwise.

The possibility of a non–monotonic relationship between variability of the environment (seen as lack of stability) and optimal length of delegation is due to the existence of two opposite effects. On the one hand, as the variability of the environment increases, there is a lower probability that an agent will fit her principal's needs for long, so it becomes more desirable for the principal to change agent frequently; this provides her with an incentive to decrease the length of the relationship. On the other hand, an increase in the variability of the environment decreases the benefit for the principal of replacing an unfit agent, since a new agent appropriately selected is less likely to remain well suited for long; this provides her with an incentive to increase the length of the relationship. [9] When the latter effect prevails, a rise in the variability of the environment leads therefore to more stable (longer) relationships.

Even though proper empirical tests are beyond the scope of the present paper, some casual evidence, mainly based on published data, seems to be consistent with the results of our analysis. The fact that the statutory provisions for the central banks of all the OECD countries establish a term length of at least 4 years for their leaders, coupled with the recognition of the need of some credibility for the management of monetary policy, seems to be consistent with the conclusion that reputation–building calls for a sufficiently long delegation arrangement; the fact that in recent decades both countries with a very unstable political environment, measured by the average duration of governments, like Italy, and a very stable one, like Canada and the U.K., all have had central bankers with quite long average length of term, while Japan, which was characterized by a degree of political stability somewhere in the middle, has had central bankers with shorter length of term, seems to be an indirect confirmation of our non–monotonicity result.[10]

The structure of the paper is the following. In section 2, we describe the

---

[9] Two similar effects have already been pointed out by Harris and Holmstrom (1987) in quite a different setting. They consider a model where the terms of the contract ruling a borrower–lender relationship depend on a randomly evolving state variable, which is observable only at a cost. The players choose when to pay such a cost and reshape the contract accordingly. Using simulations, they also obtain an U–shaped relationship between the optimal length of the agreement and the variability of the environment. In our model, we show that U–shapeness is actually not the unique possibility, and spell out mathematical conditions for both monotonicity and U–shapeness.

[10] Our calculations of the average duration of governments are based on the data in Woldendorp *et al.* (1998), while information on the average length of term of the Heads of Central banks were collected from the Banks' websites. For a little more detailed discussion of this and other sources of empirical evidence, see Section 4.

delegation game, solve it and characterize the optimal length of delegation. In section 3, we make some comparative statics. In section 4, we discuss the implications of our results in two examples of delegation and provide some empirical evidence. Section 5 concludes. All the proofs are in the Mathematical Appendices.

# 2 The model

An infinitely lived (female) principal delegates to a (male) agent the task of playing as a responder a $K$ times repeated trust game with a (gender neutral) third party, which acts as a proposer.
In a trust game, the proposer has to choose whether to *trust* ($t$) or *not* to trust ($n$) the responder. If it trusts, the responder is given the move and has to decide whether to *honor* trust ($h$) or *abuse* it ($a$). The proposer prefers to trust if the responder honors trust, and not to trust if the responder abuses trust. Whenever he is given the move, the responder prefers to behave opportunistically and abuse trust. Finally, the responder is better off when trust is given and honored than when trust is not given.

Without loss of generality, the one–shot version of the trust game can be represented as follows:

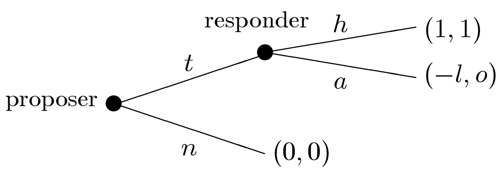


FIGURE 1. THE ONE–SHOT TRUST GAME

where $l > 0$ is the proposer's *loss* when the responder abuses its trust, while $o > 1$ is the responder's payoff when he behaves in an *opportunistic* way and abuses the proposer's trust.

We assume that the responder is either trustworthy or untrustworthy; the untrustworthy responder plays the game in FIG. 1, whereas the trustworthy always honors trust. The responder knows whether he is trustworthy, while the *a priori probability* that he is trustworthy is $p \in (0, 1)$.

The introduction of this type of incomplete information is a modelling device[11] that characterizes models à la Kreps–Wilson–Milgrom–Roberts and

[11] Hence, we do not allow the principal to devise a contract that selects trustworthy

enables to reach unique and "reasonable" conclusions in finitely repeated games. In the tradition of this literature, we assume that $p$ is small.

We also consider a second type of (real) incomplete information, which, in our opinion, captures an essential feature of delegation problems, namely the principal's concern for how the agent fits her needs: if in period $k$ the principal faces an agent of type $\theta_k$ while preferring an agent of type $\hat{\theta}_k$, she will suffer a loss $L(\theta_k, \hat{\theta}_k)$ from having an unfit agent.

We assume that the agent's type is constant through time, but the principal's loss may be not since she lives in a *changing environment*: the agent who fits her needs in period $k-1$ best may not be the best agent in period $k$. The dynamics of $\hat{\theta}_k$ is modelled as follows: in each period $k$, there is a probability $\gamma \in (0,1)$ that $\hat{\theta}_k = \hat{\theta}_{k-1}$ and a probability $1-\gamma$ that $\hat{\theta}_k$ is the realization of a random variable uniformly distributed on a circle of circumference $2d$, so that the maximum *distance* between actual and preferred agent is $d > 0$. This assumption is meant to reflect a distinction between ordinary periods, when the principal's needs stay more or less the same, and periods of structural change, when a shock affects the type of agent preferred by the principal. Thus, our model describes well situations with no trend in shocks: *ex ante*, there is no ranking of agents' types conditional on at least one shock having occurred (*i.e.*, *a priori* all agents perform equally well); moreover, shocks do not add, so the loss of having an unfit agent is bounded [12]. Finally, the above assumption makes an important distinction between two sources of variability that will be shown to have different effects on the optimal length of delegation: on the one hand, $1-\gamma$ (probability that a shock occurs) measures the lack of stability in the environment; on the other hand, $d$ (maximum distance between actual and preferred agent) measures the size (range) of the change in the environment.

We also assume that the agent's type is observable by the principal and verifiable, while the type of agent preferred by the principal in a given period is known to the principal but it is not verifiable. All types of agents are available, so in a given period the principal can hire an agent who exactly matches her needs, thus suffering no loss in that period; however, if she commits to keep the same agent for more than one period, she will risk having for some time an unfit agent (if a shock occurs).

The characteristics of the delegation arrangement between the principal

agents, because trustworthy agents actually do not exist.

[12] Assuming that the potential loss is bounded seems reasonable in most conceivable applications of our analysis: the amount of "damage" that an unfit agent can cause is limited by the value of the activity he is involved in.

and her agent may change the structure and so the outcome of the trust game between the agent and the third party. In particular, the expected time length of delegation, by affecting the perception of stability of the trust relationship, also affects the amount of trust that will arise. We model this feature by assuming that the principal chooses the length $K$ of the arrangement with her agent and, by so doing, she also chooses the number $K$ of repetitions of the trust game. $K$ is thus an endogenous variable, which is chosen by the principal taking into account both the risk (and cost) of keeping for some time an unfit agent and the payoff from the trust relationship between her agent and a third party.

The timing of the delegation game is the following:
(i) the principal ($P$) chooses which type of agent to hire and the length $K$ of the delegation arrangement;
(ii) the agent ($A$) plays a $K$ times repeated trust game as a responder with a third party ($T$), which acts as a proposer;
(iii) at the end of each stage of the trust game, there is a probability $1-\gamma$ that a shock occurs, which changes the principal's needs;
(iv) if $K$ is finite, at the end of the repeated trust game the principal hires another agent for $K$ periods, and so on.

The following figure summarizes points (i)–(iii) of the timing of the game.

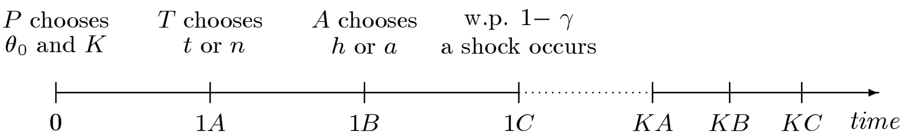


FIGURE 2. TIMING OF THE GAME

We now make three simplifying assumptions. First, the agent's ability to build a reputation for trustworthiness does not depend on his type, so the principal's problem is separable and her per–period expected utility ($\pi$) is the sum of the payoff she earns from the *trust* relationship between her agent and a third party ($\pi^T$) and from the *fitness* (or lack of fitness) of her agent ($\pi^F$). Second, the interests of the principal and the agent in the trust game are perfectly aligned; hence, we abstract from a standard feature of delegation models to focus on the length of the arrangement. Finally, $L(\theta_k, \hat{\theta}_k) = 2u\, dist(\theta_k, \hat{\theta}_k)$, where $dist(\cdot, \cdot)$ is the Euclidean distance on the circle and $u > 0$ is half the unit cost for the principal of having an *unfit* agent.

We find the sequential equilibrium of the delegation game working in two steps: first, for each length $K$ of the delegation arrangement, we write the per–period expected utility of the principal $\pi(K) = \pi^T(K) + \pi^F(K)$; then we find the length $K^*$ that maximizes $\pi(K)$.

In the first step, the principal anticipates, for each length $K$ of the delegation arrangement, the equilibrium that would obtain in the $K$ times repeated trust game between her agent and a third party. RESULT 2.1 below is a straightforward application to our setting of the Kreps and Wilson (1982) reputation–building analysis. We let $t_k$ be the probability that, in period $k$, the third party *trusts* the agent, and $h_k$ the probability that the untrustworthy agent *honors* trust.

RESULT 2.1 *Let $p_k$ be the probability the third party attaches in period $k$ to the agent being trustworthy. In the unique* [13] *sequential equilibrium,*

$$t_k = \begin{cases} 1 & \text{if } p_k > (\frac{l}{l+1})^{K-k+1} \\ \frac{o-1}{o} & \text{if } p_k = (\frac{l}{l+1})^{K-k+1} \\ 0 & \text{if } p_k < (\frac{l}{l+1})^{K-k+1} \end{cases} \tag{2.1}$$

$$h_k = \begin{cases} 1 & \text{if } p_k \geq (\frac{l}{l+1})^{K-k} \\ \frac{p_k}{1-p_k} \frac{1-(\frac{l}{l+1})^{K-k}}{(\frac{l}{l+1})^{K-k}} < 1 & \text{if } 0 < p_k < (\frac{l}{l+1})^{K-k} \\ 0 & \text{if } p_k = 0 \end{cases} \tag{2.2}$$

*where*

$$p_{k+1} = \begin{cases} p_k & \text{if } p_k > 0 \text{ and in period } k \text{ the agent is not given trust} \\ \frac{p_k}{p_k+(1-p_k)h_k} & \text{if } p_k > 0 \text{ and in period } k \text{ the agent honors trust} \\ 0 & \text{if } p_k = 0 \text{ or in period } k \text{ the agent abuses trust} \end{cases} \tag{2.3}$$

As the game unfolds, the third party updates its beliefs on the basis of its observation of the agent's behavior; as long as the latter honors trust, the probability the third party attaches to the agent being trustworthy, $p_k$, increases. Hence, the effectiveness of trust–building is affected by the number of repetitions of the game. The sequential equilibrium in RESULT 2.1 can thus be appropriately re–expressed in PROP. 2.1 as a function of the number $K$ of repetitions of the game. This is useful for solving the delegation game because it enables us to write the principal's payoff as a function of the

[13] We will henceforth rule out the zero–probability values of the parameters that give rise to multiple equilibria.

length $K$ of the arrangement with her agent. However, we also find it very interesting *per se* because it allows to rephrase reputation equilibria *à la* Kreps-Wilson-Milgrom-Roberts (and players' payoffs) as a function of the number of repetitions of the game.

PROPOSITION 2.1 *Let $\tilde{K}(p,l)$ be the integer part of $K(p,l)$, where $K(p,l) \in$ **R** is the unique solution to*

$$p = \left(\frac{l}{l+1}\right)^{K(p,l)-1}$$

*If $K < \tilde{K}(p,l)$, in the unique sequential equilibrium, $t_k = 0$ for all $k = 1, 2, \ldots, K$ (i.e., the agent is never trusted).*
*If $K \geq \tilde{K}(p,l)$, in the unique sequential equilibrium,*

$$t_k = \begin{cases} 1 & k \in \mathbb{C} \cup \mathbb{T} \\ \frac{o-1}{o} & k \notin \mathbb{C} \cup \mathbb{T} \text{ and in period } k-1 \text{ the third party gave trust and the agent honored it} \\ 0 & k \notin \mathbb{C} \cup \mathbb{T} \text{ and in period } k-1 \text{ either the third party did not give trust or the agent abused it} \end{cases} \tag{2.4}$$

$$h_k = \begin{cases} 1 & k \in \mathbb{C} \\ \frac{p}{1-p} \frac{1-(\frac{l}{l+1})^{K-1}}{(\frac{l}{l+1})^{K-1}} & k \in \mathbb{T} \\ \frac{l}{l+1} \frac{1-(\frac{l}{l+1})^{K-k}}{1-(\frac{l}{l+1})^{K-k+1}} & k \notin \mathbb{C} \cup \mathbb{T} \text{ and in the past the agent never abused trust} \\ 0 & k \notin \mathbb{C} \cup \mathbb{T} \text{ and in at least one previous period the agent abused trust} \end{cases} \tag{2.5}$$

*where $\mathbb{T} = \{K - \tilde{K}(p,l) + 1\}$ and $\mathbb{C} = \{1, 2, \ldots, K - \tilde{K}(p,l)\}$, with $\mathbb{C} = \emptyset$ if $K = \tilde{K}(p,l)$.*

Reputation equilibria *à la* Kreps-Wilson-Milgrom-Roberts, as described by PROP. 2.1, have some nice properties. First, when $K < \tilde{K}(p,l)$, $\tilde{K}(p,l)$ being the minimum number of repetitions giving rise to trust–building [14], the third

[14] In equilibrium, a reputation for trustworthiness obtains when $K \geq \tilde{K}(p,l)$, while it does not when $K < \tilde{K}(p,l)$. This seems to suggest that reputation building has a somehow dichotomous nature, since it calls for paying some type of "fixed cost" (for example, a sunk investment in time). As we will see below, this "fixed cost" gives rise to a discontinuity in the principal's expected utility function.

party does not trust the agent; so, trust cannot arise unless the parties perceive their relationship to be sufficiently stable. Second, when $K \geq \tilde{K}(p,l)$, in the first $K - \tilde{K}(p,l)$ periods the untrustworthy agent imitates the trustworthy and honors trust. Hence, in equilibrium, an increase in the number $K$ of repetitions of the game increases the number of periods belonging to the (cooperation) set $\mathbb{C}$, where trust is given and honored [15]: a more stable relationship gives rise to more trust. Finally, after the (transition) period $\mathbb{T}$, when the third party gives trust and the agent starts playing a mixed strategy, the closer the end of the game, the higher the probability that a agent who never abused trust is considered as trustworthy, so the larger the number of periods where one party honored trust, the better his reputation for trustworthiness.

Using PROP. 2.1, we can easily define the function $\pi^T(K)$. Moreover, given the dynamics of $\hat{\theta}_k$, after a shock, *a priori* all agents perform equally well, so every time the principal hires a new agent, she finds it optimal to select an agent who exactly matches her needs at that moment. As a consequence, if she keeps an agent for $K$ periods, she will start suffering a loss (with probability one) only after the first shock has occurred; this defines the function $\pi^F(K)$.

RESULT 2.2 below provides a reasonably simple expression for $\pi(K)$ when $K \geq \tilde{K}(p,l)$, while FIG. 3 illustrates the qualitative features of $\pi^T(K)$ and $\pi^F(K)$, and points out the basic trade-off between trust and fitness: an increase in $K$ rises $\pi^T(\cdot)$ but lowers $\pi^F(\cdot)$. For analytical convenience, $K$ is henceforth considered as a real number.

RESULT 2.2 *Let $K \geq \tilde{K}(p,l)$. Then*

$$\pi(K) = \pi^T(K) + \pi^F(K) = 1 - \frac{\tilde{K}(p,l) - o}{K} - ud + \frac{1}{K}\frac{1-\gamma^K}{1-\gamma}ud \qquad (2.6)$$

*where:*
*(i)* $\pi^T(K) = 1 - \frac{\tilde{K}(p,l)-o}{K}$, $\pi^T(K)$ *tends to* $1$ *as* $K$ *tends to infinite,* $\pi^{T\prime}(K) > 0$ *and* $\pi^{T\prime\prime}(K) < 0$ *for all* $K$*;*
*(ii)* $\pi^F(K) = -ud + \frac{1}{K}\frac{1-\gamma^K}{1-\gamma}ud$, $\pi^F(K)$ *tends to* $-ud$ *as* $K$ *tends to infinite,* $\pi^{F\prime}(K) < 0$ *and* $\pi^{F\prime\prime}(K) > 0$ *for all* $K$*.*

---

[15] Note that $\tilde{K}(p,l)$ is increasing in $l$ and decreasing in $p$, so the smaller the third party's loss when the agent abuses trust and the higher the probability that the agent is *trustworthy*, the larger the cooperation set.

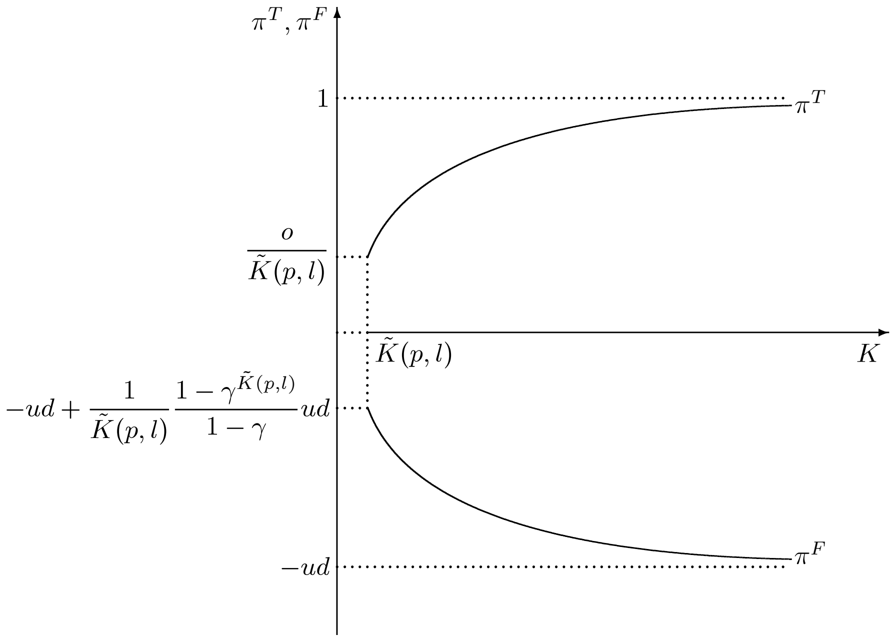


FIGURE 3. THE TRADE–OFF BETWEEN TRUST AND FITNESS

We will now provide an economic intuition for the mathematical properties of the functions $\pi^T(K)$ and $\pi^F(K)$, and give two important definitions.
In the trust game, should actions be verifiable, the parties would write a contract where the third party gives trust and the agent honors it. The principal's (cooperation) payoff would be equal to 1; during $K$ periods, the principal would thus earn $K$. In our model, however, actions are not verifiable, so if the principal hires an agent for $K$ periods, she will (only) earn $K - (\tilde{K}(p,l) - o)$ [16]. The term $\tilde{K}(p,l) - o$ can therefore be interpreted as the

[16] In the first $K - \tilde{K}(p,l)$ periods (cooperation set $\mathbb{C}$), the third party gives trust and the agent honors it, so the principal earns $K - \tilde{K}(p,l)$. In the transition period $\mathbb{T}$, the third party gives trust and the untrustworthy agent plays a mixed strategy. This implies that the agent must be indifferent between abusing trust and honoring it. Should he abuse trust, he would earn $o$ in that period and zero in all the following periods (because the third party would no longer give him trust). Hence, in the last $\tilde{K}(p,l)$ periods the principal earns (in expected terms) $o$. The overall expected utility for the principal in $K$ periods is therefore $K - \tilde{K}(p,l) + o$.

principal's *(fixed) expected cost of building reputation*: each time the principal hires a new agent for $K \geq \tilde{K}(p,l)$ periods, she "loses" $\tilde{K}(p,l) - o$ (with respect to the case where actions are verifiable) [17]. The term $\tilde{K}(p,l) - o$ plays exactly the same role as the re–contracting cost in Gray (1978), Dye (1985), Harris and Holmstrom (1987) and Cantor (1988); thus, our model endogenizes such a cost.

The larger is $K$, the less often the cost of building reputation is paid. This has a straightforward implication on the per–period principal's expected return in the repeated trust game

$$\pi^T(K) = \frac{K - (\tilde{K}(p,l) - o)}{K} = 1 - \frac{\tilde{K}(p,l) - o}{K} \tag{2.7}$$

which is increasing in $K$ because the fixed cost is split between a larger number of periods. At the limit, when $K$ tends to infinite, the effect of the per–period fixed cost vanishes, so $\pi^T(K)$ tends to the cooperation payoff 1 [18] (see RESULT 2.2 and FIG. 3).

As for

$$\begin{aligned}\pi^F(K) &= -\frac{(1-\gamma^0) + (1-\gamma) + (1-\gamma^2)\ldots + (1-\gamma^{K-1})}{K} ud \\ &= -\frac{K - \left(1 + \gamma + \gamma^2 + \ldots + \gamma^{K-1}\right)}{K} ud = -ud + \tfrac{1}{K}\tfrac{1-\gamma^K}{1-\gamma} ud\end{aligned} \tag{2.8}$$

the generic term $1 - \gamma^k$ in the numerator of the first expression represents the probability that at least one shock will occur in the first $k+1$ periods. After a shock, the agent will not match (with probability one) the principal's needs, and from that period onwards the principal will suffer (in each period) an expected loss equal to $ud$. The term $ud$ can therefore be interpreted as the principal's *(per-period) expected cost of keeping an unfit agent* [19]. $1 - \gamma^k$

---

[17] $\tilde{K}(p,l) - o$ is also the reduction in expected utility for the principal coming from the trust relationship between her agent and a third party when she hires two agents for $K \geq \tilde{K}(p,l)$ periods each rather than (only) one agent for $2K$ periods. $\tilde{K}(p,l) - o$ can therefore also be seen as the reputational cost of halving the length of the contract and interpreted as the cost for an additional agent to build a reputation for trustworthiness.

[18] This is reminiscent of a well known result by Radner (1981): in a repeated principal–agent game, the epsilon–equilibrium per–period expected utility for each player is arbitrarily close to the one–period cooperation payoff, provided that the number of repetitions is large enough.

[19] This cost positively depends on both the maximum distance between actual and required agent, $d$ (*i.e.*, the size of the possible changes in the environment), and half the unit loss of having an unfit agent, $u$.

is increasing in $k$. Hence, the larger is $K$, the more often the above cost is paid. This explains why $\pi^F(K)$ is decreasing in $K$. At the limit, when $K$ tends to infinite, the agent (almost) always does not fit his principal's needs, so $\pi^F(K)$ tends to $-ud$ (see RESULT 2.2 and FIG. 3).

RESULT 2.2 provides a mathematical expression for $\pi(K)$ when $K \geq \tilde{K}(p,l)$. To complete the first step of the solution, we also need to write $\pi(K)$ for $K < \tilde{K}(p,l)$. According to PROP. 2.1, in the repeated trust game the third party never trusts the agent, so

$$\pi(K) = \pi^F(K) = -ud + \frac{1}{K}\frac{1-\gamma^K}{1-\gamma}ud \tag{2.9}$$

Let us finally analyze the second step of the solution procedure: the principal chooses the (optimal) length $K^*$ that maximizes $\pi(K)$. To ease the interpretation of the results, we treat it as a new two–stage procedure: first, the principal chooses whether to let her agent build a reputation for trustworthiness by setting $K \geq \tilde{K}(p,l)$ or renounce reputation–building and set $K < \tilde{K}(p,l)$; then, for each of the two cases, the principal chooses the length that maximizes her per–period expected utility ($K^R$ and $K^N$ will henceforth denote the optimal length of the delegation arrangement for the principal when, in the first stage, she chose to let her agent, respectively, build a *reputation* for trustworthiness and *not* build it).

If in the first stage the principal chose to let her agent build a reputation for trustworthiness, in the second stage she will maximize EQ. (2.6). PROP. 2.2 below characterizes the unique candidate to an interior solution, $\bar{K}$, and shows that $K^R = \bar{K}$ whenever $\bar{K}$ exists *and* $\bar{K} \geq \tilde{K}(p,l)$. On the other hand, when either $\bar{K}$ does not exist or $\bar{K} < \tilde{K}(p,l)$, a boundary solution obtains (respectively, $K^R = \infty$ and $K^R = \tilde{K}(p,l)$). As for the case where, in the first stage, the principal chose to renounce reputation–building, in the second stage she will maximize EQ. (2.9). $\pi^F(K)$ is strictly decreasing in $K$, so $K^N = 1$.

Given the solution of the second stage, the principal solves the first stage by comparing $\pi(K^R)$ and $\pi(K^N) = \pi(1) = 0$. PROP. 2.2 performs such a comparison and characterizes the optimal length $K^* \in \{K^N, K^R\}$ of the delegation arrangement.

PROPOSITION 2.2 *Let*

$$\beta = \frac{\tilde{K}(p,l) - o}{ud} \tag{2.10}$$

$$g(K) = \gamma^K(1 - K\ln\gamma) + \beta - \gamma\beta - 1 \tag{2.11}$$

*and*

$$\bar{K} \textit{ be such that } g(\bar{K}) = 0 \tag{2.12}$$

*(i) If* $\gamma > 1 - \dfrac{1}{\beta}$, $\bar{K}$ *exists and is unique.*

*For* $K \geq \tilde{K}(p,l)$, $\pi(K)$ *reaches its maximum at* $K^R = \max\{\tilde{K}(p,l), \bar{K}\}$. *The optimal length of delegation is* [20]

$$K^* = \begin{cases} K^R = \max\{\tilde{K}(p,l), \bar{K}\} & \textit{if } \pi(\max\{\tilde{K}(p,l), \bar{K}\}) \geq 0 \\ K^N = 1 & \textit{if } \pi(\max\{\tilde{K}(p,l), \bar{K}\}) < 0 \end{cases}$$

*(ii) If* $\gamma \leq 1 - \dfrac{1}{\beta}$, $\bar{K}$ *does not exist.*

*For* $K \geq \tilde{K}(p,l)$, $\pi(K)$ *is strictly increasing in* $K$, *so* $K^R = \infty$. *The optimal length of delegation is*

$$K^* = \begin{cases} K^N = 1 & \textit{if } ud > 1 \\ K^R = \infty & \textit{if } ud \leq 1 \end{cases}$$

Prop. 2.2 shows that there are four possible types of equilibria. Three of them are boundary solutions. First, when $K^* = \infty$, the principal accepts to have almost always an unfit agent, thus suffering a per–period expected loss equal to $ud$, in order to induce the third party to always trust her agent, thus earning 1 in each period. The (overall) per–period expected payoff for the principal is therefore $1 - ud$. Second, when $K^* = 1$, the principal always has the most suitable agent, so she suffers no loss from fitness, but has to renounce reputation–building. The (overall) per–period expected payoff for the principal is therefore 0 [21]. Third, when $K^* = \tilde{K}(p,l)$, the principal selects the minimum length for the contract that induces the third party to trust her agent. In addition to the three boundary solutions, there may also be an internal solution $K^* = \bar{K} \in (\tilde{K}(p,l), \infty)$. The dependence of the equilibria on the parameters of the model will be analyzed in the next section.

[20] In the zero–probability event where the principal is indifferent between letting her agent build a reputation for trustworthiness (but risking keeping for some time an unfit agent) and renouncing reputation–building (but always having the best agent), we assume that she chooses the first alternative.

[21] It is worth noting that, if $ud < 1$, then $K = \infty$ is better than $K = 1$, so $K^* = K^R$, *i.e.*, the principal always deems it optimal to let her agent build a reputation for trustworthiness.

# 3 Comparative statics

According to PROP. 2.2, the optimal length of the delegation arrangement depends on many parameters: $p$ and $l$ (which affect $\tilde{K}(p,l)$), $o$, $u$, $d$ and $\gamma$. For simplicity, we have chosen to focus on three functions of such parameters: the principal's *per-period expected cost of keeping an unfit agent* ($ud$), the principal's *(fixed) expected cost of building reputation* ($\tilde{K}(p,l)-o$) and the *per-period probability that a shock will not occur* ($\gamma$).
An increase in $ud$ makes it more costly for the principal to keep the same agent for more periods (larger loss if a shock occurs). An increase in $\tilde{K}(p,l)-o$ makes it more costly for the principal to terminate a relationship with an agent and start a relationship with a new agent who is called to build a reputation for trustworthiness (larger fixed cost every $K$ periods). Finally, an increase in $\gamma$ makes it less costly for the principal to keep the same agent for more periods (lower probability that a shock will occur). Thus, one might conjecture that the optimal length of delegation is *(i)* decreasing in $ud$, *(ii)* increasing in $\tilde{K}(p,l)-o$, *(iii)* increasing in $\gamma$.
We will now show that *(i)* is unconditionally true, *(ii)* is true only if the change in $\tilde{K}(p,l)-o$ does not affect the principal's decision to let her agent build a reputation for trustworthiness, *(iii)* may be false because an increase in $\gamma$ also makes it more profitable for the principal to hire a new agent (higher expected number of periods where the principal suffers no loss), so a change in $\gamma$ gives rise to two opposite effects, and the optimal length of delegation is either increasing or decreasing in $\gamma$ according to which of the two effects prevails.

## 3.1 Costs and benefits of a stable relationship

In this subsection, we study how the optimal length $K^*$ of delegation is affected by a change in either the principal's *per-period expected cost of keeping an unfit agent* ($ud$) or the principal's *(fixed) expected cost of building reputation* ($\tilde{K}(p,l)-o$).

The following proposition shows how $K^*$ is affected by a change in $ud$.

PROPOSITION 3.1 *There exists $\overline{\varphi}>0$ such that*

$$K^* = \begin{cases} K^R \geq \tilde{K}(p,l) \ \textit{decreasing in } ud & \textit{if } ud \leq \overline{\varphi} \\ K^N = 1 & \textit{if } ud > \overline{\varphi} \end{cases}$$

The optimal length of delegation is decreasing [22] in $ud$. As long as $ud \leq \overline{\varphi}$,

[22] The terms *decreasing* and *increasing* will henceforth be used as synonyms for *non–*

the principal deems it optimal to let her agent build a reputation for trustworthiness ($K^* = K^R$) and reacts to an increase in the per–period expected cost of having an unfit agent by decreasing the length of the delegation arrangement, thus lowering the probability of paying such a cost. However, when $ud$ exceeds the threshold level $\overline{\varphi}$, having an unfit agent is so costly that the principal chooses to renounce reputation–building ($K^* = K^N$): the optimal length of delegation jumps to one, and in each period the principal hires an agent who exactly matches her needs.

Proposition 3.2
*(i) If $ud < 1$, $K^* = K^R \geq \tilde{K}(p,l)$ increasing in $\tilde{K}(p,l) - o$.*
*(ii) If $ud \geq 1$, there exists $\bar{\psi}$ such that* [23]

$$K^* = \begin{cases} K^R = \max\{\tilde{K}(p,l), \bar{K}\} \\ \textit{increasing in } \tilde{K}(p,l) - o & \textit{if } \tilde{K}(p,l) - o \leq \bar{\psi} \\ \\ K^N = 1 & \textit{if } \tilde{K}(p,l) - o > \bar{\psi} \end{cases}$$

An increase in the (fixed) expected cost of building reputation $\tilde{K}(p,l) - o$ has two effects. On the one hand, it reduces the incentive for the principal to let her agent build a reputation for trustworthiness. On the other hand, when the principal chooses to let her agent build a reputation for trustworthiness, it provides the principal with an incentive to increase the length of the delegation arrangement in order to pay less often such a fixed cost.

When $ud \geq 1$, the principal chooses to let her agent build a reputation for trustworthiness ($K^* = K^R$) only when the (fixed) expected cost of building reputation is sufficiently low ($\tilde{K}(p,l) - o \leq \bar{\psi}$). In that case, the principal reacts to an increase in the (fixed) expected cost of building reputation by increasing the length of the delegation arrangement, thus paying such a cost less often. However, when such a cost exceeds the threshold level $\bar{\psi}$, the principal chooses to renounce reputation-building ($K^* = K^N$), so the optimal length of delegation jumps to one, and in each period the principal hires an agent who exactly matches her needs. Finally, when $ud < 1$, the principal always deems it optimal to let her agent build a reputation for trustworthiness ($K^* = K^R$), so, as before, she reacts to an increase in the

---

*increasing* and *non–decreasing*. Indeed, in a boundary solution ($\tilde{K}(p,l)$ or $\infty$) $K^R$ may be insensitive to the value of some parameters within a given interval.

[23] We rule out the (not particularly interesting) case where, for a given $\tilde{K}(p,l)$, $K^* = K^N = 1$ for all $o \leq \tilde{K}(p,l)$ or, for a given $o$, $K^* = K^N = 1$ for all $\tilde{K}(p,l) \geq o$, so that the optimal length of the arrangement does not depend on $\tilde{K}(p,l) - o$.

(fixed) expected cost of building reputation by increasing the length of the delegation arrangement, thus paying such a cost less often.

As a conclusion, the comparative statics analysis of this section confirms the intuition on the effects of the two different types of costs on the optimal length of delegation, provided the principal decides to let her agent build a reputation for trustworthiness. More interestingly, increases in both these costs may trigger a regime switch, where the principal chooses to give up reputation-building, and the optimal length of delegation jumps to one.

## 3.2 Stability of the environment

In this subsection, we study how the optimal length $K^*$ of delegation is affected by changes in the stability of the environment as measured by the *per-period probability that a shock will not occur* ($\gamma$).

For some values of the parameters, $K^*$ is equal to $\bar{K}$, which depends on $\gamma$ (EQS. (2.11) and (2.12)). The following result characterizes the behavior of the function $\bar{K}(\gamma)$ at the boundaries of its domain [24].

RESULT 3.1
*(a) When $\beta \geq 1$, as $\gamma$ tends to either $1 - \frac{1}{\beta}$ or 1, $\bar{K}(\gamma)$ tends to infinite;*
*(b) When $\beta < 1$, as $\gamma$ tends to 0, $\bar{K}(\gamma)$ tends to zero; as $\gamma$ tends to 1, $\bar{K}(\gamma)$ tends to infinite.*

RESULT 3.1 implies that when $\beta \geq 1$, there exist $\gamma_1, \gamma_2 \in (1 - \frac{1}{\beta}, 1)$ such that the function $\bar{K}(\gamma)$ is strictly decreasing in $\gamma$ for all $\gamma \in (1 - \frac{1}{\beta}, \gamma_1)$ and strictly increasing in $\gamma$ for all $\gamma \in (\gamma_2, 1)$. Hence, when $\beta \geq 1$, the function $\bar{K}(\gamma)$ is always non–monotonic in $\gamma$.

Unfortunately, we have been unable to fully investigate analytically the behavior of the function $\bar{K}(\gamma)$ in the interior of its domain. However, numerical simulations unequivocally suggest that $\gamma_1 = \gamma_2$, so $\bar{K}(\gamma)$ is strictly decreasing in $\gamma$ for "low" values of $\gamma$ and strictly increasing in $\gamma$ for "high" values of $\gamma$ [25]. We confidently write this "conjecture" as a formal result [26].

---

[24] From PROP. 2.2, when $\beta \geq 1$, $\bar{K}(\gamma)$ exists if (and only if) $\gamma \in (1 - \frac{1}{\beta}, 1)$; when $\beta < 1$, $\bar{K}(\gamma)$ exists for all $\gamma \in (0, 1)$.

[25] For a somehow similar result in quite a different model, see Harris and Holmstrom (1987). This might suggest that the above result is robust to (even) drastic changes in the model. However, we show below that when $\beta < 1$, $\bar{K}(\gamma)$ is increasing in $\gamma$, so U–shapeness seems to be a less general result than one could expect after reading Harris and Holmstrom (1987).

[26] We do not delve into the methodological issue of whether and when numerical simu-

RESULT 3.2 *When $\beta \geq 1$, there exists $\bar{\gamma} \in (1 - \frac{1}{\beta}, 1)$ such that*

$$\begin{cases} d\bar{K}(\gamma)/d\gamma < 0 & \textit{for all } \gamma \in (1 - \frac{1}{\beta}, \bar{\gamma}) \\ d\bar{K}(\gamma)/d\gamma = 0 & \textit{for } \gamma = \bar{\gamma} \\ d\bar{K}(\gamma)/d\gamma > 0 & \textit{for all } \gamma \in (\bar{\gamma}, 1) \end{cases}$$

For example, FIG. 4 shows the function $\bar{K}(\gamma)$ when $\beta$ is, respectively, 1.20, 1.50 and 1.80 (the higher the value of $\beta$, the higher the curve).

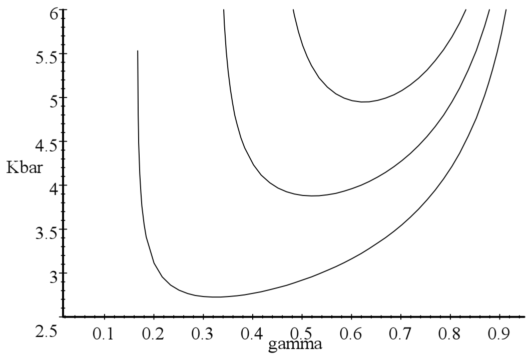


FIGURE 4: When $\beta \geq 1$, $\bar{K}(\gamma)$ is not monotonic in $\gamma$

The following proposition also considers boundary solutions, and shows how $K^*$ is affected by a change in $\gamma$ when $\beta \geq 1$.

PROPOSITION 3.3 *Let $\beta \geq 1$.*
*(i) If $ud < 1$,*

$$K^* = \begin{cases} K^R = \infty & \gamma \in (0, 1 - \frac{1}{\beta}] \\ K^R = \max\{\tilde{K}(p,l), \bar{K}(\gamma)\} \textit{ decreasing in } \gamma & \gamma \in (1 - \frac{1}{\beta}, \bar{\gamma}) \\ K^R = \max\{\tilde{K}(p,l), \bar{K}(\gamma)\} \textit{ increasing in } \gamma & \gamma \in [\bar{\gamma}, 1) \end{cases}$$

*(ii) If $ud \geq 1$, there exists $\gamma^R \in (1 - \frac{1}{\beta}, 1)$ such that* $\pi(\max\{\tilde{K}(p,l), \bar{K}(\gamma^R)\}) =$

---

lations can actually *prove analytically* an assertion. However, we think that, *in practical terms*, checking the validity of the following proposition for all $\beta \in (\mathbf{0}, \mathbf{10})$ (considering changes by $\mathbf{0.10}$) and for all $\gamma \in (\mathbf{0}, \mathbf{1})$ (considering changes by $\mathbf{0.01}$) is more than enough for attaining an absolute confidence on its correctness.

0, *and*

$$K^* = \begin{cases} K^N = 1 & \gamma \in (0, \gamma^R) \\ K^R = \max\{\tilde{K}(p,l), \bar{K}(\gamma)\} \text{ decreasing in } \gamma & \gamma \in [\gamma^R, \max\{\gamma^R, \bar{\gamma}\}) \\ K^R = \max\{\tilde{K}(p,l), \bar{K}(\gamma)\} \text{ increasing in } \gamma & \gamma \in [\max\{\gamma^R, \bar{\gamma}\}, 1) \end{cases}$$

We now provide a possible intuition for the above results in terms of the two effects mentioned in SEC. 3. Let us initially consider the case where $ud < 1$, so that the principal always deems it optimal to let her agent build a reputation for trustworthiness ($K^* = K^R$). When $\gamma$ is "very low" ($\gamma \leq 1 - \frac{1}{\beta}$), the benefit from changing the agent is relatively small and not lasting, so the principal completely disregards such a possibility and selects the longest possible arrangement ($K^* = \infty$). As the environment becomes more stable but $\gamma$ is still "rather low" ($\gamma \in (1 - \frac{1}{\beta}, \bar{\gamma})$), the principal can reason as if she could disregard what happens after few periods, since from then on the agent will "almost certainly" be unfit for the job. The only really meaningful effect of an increase in $\gamma$ is that it makes it more likely that in the first few periods following a replacement the new agent will still be fit for the job (*i.e.*, it increases the value of hiring a new agent) [27]. Hence, the principal chooses to benefit more often from those large initial payoffs by shortening the length of the delegation arrangement ($K^* = \max\{\tilde{K}(p,l), \bar{K}(\gamma)\}$ decreasing in $\gamma$). Finally, when $\gamma$ is "high" ($\gamma \in [\bar{\gamma}, 1)$), the probability that an agent will still be fit for the job after many periods is sufficiently high that the principal cannot disregard what happens then. An increase in $\gamma$ makes the fitness of the agent last for more periods (in expected terms); this effect prevails [28], so it becomes more desirable for the principal to select longer term arrangements ($K^R = \max\{\tilde{K}(p,l), \bar{K}(\gamma)\}$ increasing in $\gamma$). As for the case where $ud \geq 1$, the only meaningful difference is that in a relatively unstable environment ($\gamma < \gamma^R$) the principal finds it too costly to let her agent build a reputation for trustworthiness, since this would entail a high probability of keeping for some time an unfit agent, so she deems it preferable to renounce reputation–building: the optimal length of the arrangement is $K^* = K^N = 1$, and in

---

[27] This apparently speculative argument is somehow supported by the following analytical result. If let $\Pi^F_k(\gamma) = -(1 - \gamma^{k-1})ud$ be the principal's expected payoff from his agent's type in the $k^{th}$ period of his activity, then $\gamma < \frac{k-1}{k} \Longrightarrow \frac{\partial \Pi^F_k(\gamma)}{\partial \gamma} > \frac{\partial \Pi^F_{k+1}(\gamma)}{\partial \gamma} > \frac{\partial \Pi^F_{k+2}(\gamma)}{\partial \gamma} > \ldots$, *i.e.*, when $\gamma$ is "low", an increase in $\gamma$ benefits more the first periods following a replacement.

[28] This is somehow the word version of the mathematical result $\gamma > \frac{k-1}{k} \Longrightarrow \frac{\partial \Pi^F_{k+1}(\gamma)}{\partial \gamma} > \frac{\partial \Pi^F_k(\gamma)}{\partial \gamma} > \frac{\partial \Pi^F_{k-1}(\gamma)}{\partial \gamma} > \ldots$

each period the principal hires an agent who exactly matches her needs.

RESULT 3.1 also implies that when $\beta < 1$, there exist $\gamma_3, \gamma_4 \in (0,1)$ such that the function $\bar{K}(\gamma)$ is strictly increasing in $\gamma$ for all $\gamma \in (0, \gamma_3) \cup (\gamma_4, 1)$ (and so it cannot be U–shaped). Straightforward numerical simulations suggest that $\gamma_3 = \gamma_4$, so the function $\bar{K}(\gamma)$ is strictly increasing in $\gamma$. Once again, we write this "conjecture" as a formal result.

RESULT 3.3 *If* $\beta < 1$, $d\bar{K}(\gamma)/d\gamma > 0$ *for all* $\gamma$.

For example, FIG. 5 shows the function $\bar{K}(\gamma)$ when $\beta$ is, respectively, 0.20, 0.50 and 0.80 (again, the higher the value of $\beta$, the higher the curve).

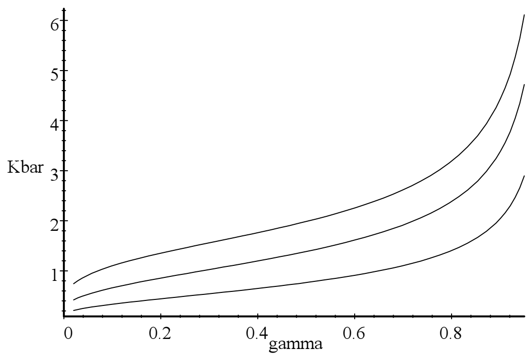


FIGURE 5: When $\beta < 1$, $\bar{K}(\gamma)$ is strictly increasing in $\gamma$

The following proposition also considers boundary solutions and shows how $K^*$ is affected by a change in $\gamma$ when $\beta < 1$.

PROPOSITION 3.4 *Let* $\beta < 1$.
*(i) If* $ud < 1$, *there exists* $\gamma^I \in (0,1)$ *such that* $\bar{K}(\gamma^I) = \tilde{K}(p,l)$, *and*

$$K^* = \begin{cases} K^R = \tilde{K}(p,l) & \gamma \in (0, \gamma^I) \\ K^R = \bar{K}(\gamma) \text{ increasing in } \gamma & \gamma \in [\gamma^I, 1) \end{cases};$$

*(ii) If* $ud \geq 1$, *there exists* $\gamma^R \in (1 - \frac{1}{\beta}, 1)$ *such that* $\pi(\max\{\tilde{K}(p,l), \bar{K}(\gamma^R)\}) = 0$, *and*

$$K^* = \begin{cases} K^N = 1 & \gamma \in (0, \gamma^R) \\ K^R = \max\{\tilde{K}(p,l), \bar{K}(\gamma)\} \text{ increasing in } \gamma & \gamma \in [\gamma^R, 1) \end{cases}$$

Let us initially consider the case where $ud < 1$, so that the principal always deems it optimal to let her agent build a reputation for trustworthiness ($K^* = K^R$). When $\gamma$ is "very low", it is "very likely" that the agent will soon fail to meet his principal's needs; furthermore, when $\beta < 1$, the cost for the principal of keeping an unfit agent is high relative to the cost of building reputation: to have the "right" agent even only for one period is too important, so $K^* = \tilde{K}(p,l)$. As the environment gets more stable, the "right" agent stays "right" for longer, so the principal pays more attention to reputation by increasing the length of the arrangement ($K^R = \bar{K}(\gamma)$ increasing in $\gamma$). As for the case where $ud \geq 1$, the most important difference is that in a relatively unstable environment ($\gamma < \gamma^R$) the principal finds it too costly to let her agent build a reputation for trustworthiness, so $K^* = K^N = 1$, and in each period the principal hires an agent who exactly matches her needs.

COR. 1 follows immediately from PROPP. 3.3 and 3.4.

COROLLARY 1 *If* $\gamma > \max\left\{1 - \frac{1}{\beta}, \gamma^R\right\}$, *then* $K^* = 1, \infty$ *is increasing in* $\gamma$ *if* $\beta < 1$, *i.e., if* $\tilde{K}(p,l) - o < ud$, *while it is U–shaped otherwise.*

From PROP. 3.1 and COR. 1, we have that while the optimal length of delegation is decreasing in the variability of the environment when such variability is interpreted as the size (width) of the change in the environment ($d$, for a given $u$, provided the principal chooses to let her agent build a reputation for trustworthiness), this need not be true when it is interpreted as lack of stability ($1 - \gamma$). Thus, the two alternative measures of variability considered in this paper have different effects on the optimal length of delegation.

# 4 Discussion

Our model is meant to be a stylized representation of many multi–task delegation settings where moral hazard games are played in changing environments.

In this section, we consider two general sets of applications, but mainly focus on one specific example for each of them, which we had already referred to in the introduction: the delegation of monetary policy to independent central bankers and the delegation of authority to managers within firms. We first show how these two examples fit the model, then discuss some implications of our results for those settings and finally provide some evidence based on published data that seems to be consistent with the results of our analysis.[29]

Our first general set of applications are public officials entrusted with discretionary power that may be affected by time–inconsistency, such as elected representatives in the executive branch of government or appointed members of independent agencies like central banks or antitrust authorities. A third party, be it a foreign government, an interest group, the private sector or any other subject interacting with the public administration, has to decide whether to trust the official. The official needs a sufficiently long mandate to build a reputation for trustworthiness, but he may become, as time passes, unsuitable to represent the preferences of his principal, be it the whole electorate or a branch of the public administration.

To fix ideas, we consider the example of delegation of monetary policy to an independent central banker. The private sector has to decide whether to trust ($t$) the central banker, *i.e.* take it for granted that the announced inflation target will actually be pursued, or not to trust him ($n$) and expect an inflation bias; the central banker selects the monetary policy either to keep inflation close to the target ($h$) or to accommodate the government's expansionary objectives ($a$). In the case under consideration, $l$ and $o$ represent, respectively, the loss for the private sector in terms of reduced purchasing power caused by unexpected inflation and the gain for the government (measured by nominal output boost and/or "feel good" effect in the electoral cycle) when the central banker selects a loose monetary policy; 1 is both the private sector's and the government's payoff from low inflation; $\gamma$ is a measure of stability of the government's preferences on some other dimension of the central banker duties, such as the prudential supervision of credit institutions and the pursuit of stability of the financial system, while $ud$ is the expected loss of a change of tastes in that respect. Hence, the variability of the environment in this setting can be interpreted as a way of modelling the potential effect of the political cycle on the management of monetary policy. Indeed, the change in the government's preferences can be thought as the consequence of a general election or of some other major change in the political scenario: for instance, a newly elected government may prefer a less interventionist central banker on mergers and acquisition in the credit sector.

Our second general set of applications concerns the delegation of authority within firms from owners or higher levels of the hierarchy to managers or workers. While many responsibilities within an organization entail at least a

[29] Most of the evidence refers to central bankers, because sufficiently detailed data on personal relationships within firms would require a micro–level empirical investigation that is definitely beyond the scope of the present work.

certain amount of trust building, the skills required by some complementary duty may change over time following changes in the operating environment. As a more specific example, we consider a manager to whom the owner of a firm delegates both the supervision of the employees and some strategic decisions on products' development. The employees have to choose whether to trust ($t$) the manager and work hard, expecting the manager to reward them with a bonus ($h$), or to shirk ($n$), fearing that the manager will not pay any bonus ($a$). In the case under consideration, $l$ represents the cost of effort for the employee, $1 + l$ the value of the (promised) bonus and $o$ the owner's gain from benefiting from a high level of effort without paying any bonus; $\gamma$ is a measure of stability of the environment where the firm operates and $ud$ is the expected loss from a change in product market conditions, which modifies the optimal features required in a product developing manager (*e.g.,* the launch of new products becomes more profitable than incremental innovations on old ones).

Two features of our model have interesting implications for the above, as well as for many other examples of delegation of activities requiring trust building. First, the cost of building reputation is indivisible and does not change with the length of the delegation arrangement. This implies that there is a minimum length for the delegation arrangement that is compatible with trust being built. Second, the optimal length of delegation increases as the cost of building reputation gets larger.[30] Hence, central bankers' mandates length should always exceed a threshold; moreover, the more costly for the private sector are inflation surprises, the longer should be the central bankers' mandates. In a similar way, managers' positions should be more stable where motivating workers is more difficult.

Some observations on the length of central bankers's mandates in major world economies are consistent with our predictions. The statutory provisions for the central banks of the OECD countries establish a term length of at least 4 years for their leaders, thus providing them with a sufficiently long horizon for reputation–building. Moreover, during the period between 1960 and the beginning of the present century, the average lengths of central bankers's mandates of the G7 countries were positively correlated with the average rates of unemployment for the same countries, which we could interpret as a measure of the cost of building reputation.[31] Canada and

[30] Quite intuitively, this cost is increasing in $l$, *i.e.*, the amount of the loss for the third party when trust is abused; less obviously, it is decreasing in $o$, *i.e.*, the gain from abusing trust (the larger such a gain, the stronger the signal of trustworthiness conveyed by the act of honoring trust, and so the more quickly reputation is built).

[31] We think that it is quite reasonable to assume that the larger is the rate of unem-

Italy, which were characterized by comparatively high average unemployment rates (7.6% and 7.5% respectively), also showed longer average lengths of terms of their central banker (more than 10 years), compared with an average of 9 years and 5.9% of unemployment in the U.S, 6.7 years and 4.2% of unemployment in Germany, 5 years and 2.2% of unemployment in Japan. [32]

Our analysis also shows that two different dimensions of the variability of the environment (the size and the likelihood of potential changes) affect the optimal stability of delegation arrangements in different ways. First, larger potential changes in the environment are always associated to less stable delegation. Hence, we should observe shorter mandates for public officials in countries where the political spectrum is more widely polarized, and more stable positions for occupations within firms' hierarchies where reputation–building duties are more important than the technical competence. On the first respect, we have compared the classic measure of political polarization by Laver and Hunt (1992) [33] with the average length of term of central bankers in the G7 countries after the World War II: Japan, with a comparatively high degree of polarization (5.91), has a relatively short average length of 5.18 years; at the other extreme of the spectrum, the US have a low degree of polarization (3.9) and a long average length of term (9.5); finally, Germany is somewhere in the middle, with a degree of polarization of 4.27 and an average length of 6.71. Some indirect evidence could also be collected on the stability of jobs by examining data on employee tenure in different occupations for the U.S. as provided by the Bureau of Labor Statistics. [34] In all the four reported surveys, conducted every two years between 2000 and 2006, within the highest–skilled category of "Management, professional and related occupations" the median years of tenure for management, where

ployment, the larger is the optimal size of an inflation surprise for the government, but then the more costly is for the private sector being fooled, and so the larger is the cost of building reputation for the government.

[32] The average rates of unemployment are simple averages over the period 1960-2001 of the data provided by the Labour Force Survey of the AMECO Dataset by the European Commission and extracted on September 16, 2007. We based our calculations of the average length of terms of the heads of the central banks for the corresponding period on the information provided on the banks' websites.

[33] Among the different dimensions along which polarization of political parties' attitudes has been measured by Laver and Hunt (1992), we chose to use the issue of "public ownership of business and industry", because it seemed to be the one that most appropriately reflects differences in opinion about central bankers' activities. The degree of polarization we use is calculated as the standard deviation of the distribution of measures of political parties' attitude as recorded in surveys among political scientists.

[34] U.S. Department of Labor (2006).

reputation is presumably a crucial asset, were at least 20% higher than those for professional and related occupations, where the relevance of specific skills seems more important.[35]

We also showed that the effect of a change in the likelihood that a structural break occurs, *i.e.,* in the stability of the environment, is not unequivocal and gives rise to a non–monotonicity. When stability decreases, adjustments to changes become less profitable, but at the same time more frequently needed: if the issue of reputation building is not overshadowed by the concern for the fitness of the delegate ($ud < \tilde{K}(p,l) - o$), the former effect prevails when the environment is very unstable, while the latter does in relatively stable settings. On the other hand, when the concern for the fitness is foremost ($ud > \tilde{K}(p,l) - o$), the (more intuitive) increasing relation between stability and length is re–established.

Reputation is crucial for the activity of most public officials, so we expect their mandates to be long both in countries with very stable and very unstable political systems. Indeed, in the case of central bankers, the correlation between the length of mandates for the G7 countries and the stability of the political systems as measured by the average duration of governments shows a non–monotonicity.[36] Italy and Canada have the longest average length of mandates for their central bankers (9.8 and 10 years respectively), but while Italy has the shortest average duration of governments (328 days in the period 1946–1995), Canada has one of the longest (927 days in the period 1945–1993); between the two extremes are Japan (with a 5 years average length of mandates and a 465 days average duration of governments in the period 1946–1996) and Germany (with a 7.2 years average length of mandates and a 660 days average duration of governments during the period 1949–1994).

With respect to the stability of the environment, some indirect evidence could also be collected on employee tenure by examining U.S. data concerning different sectors. [37] The level of aggregation of the data does not allow to single out different occupations within each sector: jobs where reputation–building is important are in the same pool with a majority of jobs where this is not the case. We think that, on the whole, the aggregate data is likely to represent a situation where the concern for fitness is the driving force

---

[35] The median years of tenure were 5.3 for management vs. 4.4 for professionals in 2000, 5.6 vs. 4.2 in 2002, 6.0 vs. 4.7 in 2004 and 6.0 vs. 5.0 in 2006.

[36] Our calculation of the average duration of governments is based on the data provided by Woldendorp *et al.* (1998), which do not consider the U.S.. This is the reason why we also excluded the U.S. from our analysis of correlation.

[37] U.S. Department of Labor (2006).

in determining the length of tenure, so we expect to find a monotonic relation between stability and length of tenure. The data seem to be consistent with such an expectation. Indeed, the median years of tenure in 2006 were higher in sectors that (at least we conjecture) have less variable demand. For example, the median tenure in manufacturing (5.5) was higher than in construction (3.0), which in turn exceeded the one in leisure and hospitality (1.9). At a more disaggregated level, within the trade sector, the median tenure in wholesale was 4.6 as opposed to 2.8 in retail; within the information sector, it was 5.3 in publishing as opposed to 1.9 in motion picture and sound recording. Finally, at a more aggregated level, the median tenure for the public sector (6.9) was almost double the one of the private sector (3.6).

# 5 Conclusion

We have shown that the Kreps–Wilson–Milgrom–Roberts framework for modelling reputation–building can fruitfully be employed to study the optimal length of delegation and to endogenize the benefits of long–term delegation arrangements through the recognition of the value of stability for trust. The optimal delegation will generally be longer the higher is the cost of building reputation and the lower is the cost of having an unfit agent in the delegated activity.

As for the effect of the degree of stability of the environment where the delegated activity is carried out, while one would expect that long delegation only obtains in very stable environments, our analysis showed that when the delegate's specific skills are not crucial, a rise in instability of the environment may lead to a longer duration in the relationship between principal and delegate, so we should expect, and not seldom observe, as our title says, "stable delegation in an unstable environment".

## References

AGHION, P., AND J. TIROLE (1997): "Formal and Real Authority in Organizations," *Journal of Political Economy*, 105, 1–29.

ANDERHUB, V., D. ENGELMANN, AND W. GUTH (2002): "An experimental study of the repeated trust game with incomplete information," *Journal of Economic Behavior & Organization*, 48, 197–216.

BACKUS, D., AND J. DRIFFILL (1985): "Inflation and Reputation," *American Economic Review*, 75(3), 530–538.

BARRO, R. J. (1986): "Reputation in a Model of Monetary Policy with Incomplete Information," *Journal of Monetary Economics*, 17, 3–20.

BARRO, R. J., AND D. GORDON (1983): "Rules, discretion and reputation in a model of monetary policy," *Journal of Monetary Economics*, pp. 101–121.

BERG, J., J. DICKHAUT, AND K. MCCABE (1995): "Trust, Reciprocity, and Social History," *Games and Economic Behavior*, 10, 122–142.

BRANDT, J., AND N. FIGUERAS (2003): "An Exploration of Reputation Formation in Experimental Games," *Journal of Economic Behavior & Organization*, 50, 89–115.

BULL, C. (1987): "The Existence of Self–Enforcing Implicit Contracts," *Quarterly Journal of Economics*, 102(1), 147–159.

BURNHAM, T., K. MCCABE, AND V. SMITH (2000): "Friend–or–foe intentionality priming in an extensive form trust game," *Journal of Economic Behavior & Organization*, 43, 57–73.

CAMERER, C. F. (2003): *Behavioral Game Theory: Experiments in Strategic Interaction.* Princeton University Press, Princeton, New Jersey.

CANTOR, R. (1988): "Work Effort and Contract Length," *Economica*, 55, 343–353.

CHRISTOFIDES, L. N., AND D. A. WILTON (1983): "The Determinants of Contract Length: An Empirical Analysis Based on Canadian Micro Data," *Journal of Monetary Economics*, 12, 309–319.

COLOMBO, F., AND G. MERZONI (2004): "Reputazione, flessibilità e durata ottima dei contratti," *Economia politica*, 21(2), 233–268.

——— (2006): "In Praise of Rigidity. The bright side of long–term contracts in repeated trust games," *Journal of Economic Behavior & Organization*, 59(3), 349–373.

DANZIGER, L. (1988): "Real shocks, efficient risk sharing, and the duration of labor contracts," *Quarterly Journal of Economics*, pp. 435–440.

DYE, R. A. (1985): "Optimal length of labor contracts," *International Economic Review*, 26(1), 251–270.

EUROPEAN COMMISSION (2007): "Labour Force Survey," *AMECO Dataset. Data extracted on 16.09.2007.*

FERSHTMAN, C., AND K. L. JUDD (1987): "Equilibrium Incentives in Oligopoly," *American Economic Review*, 77(5), 927–940.

FUDENBERG, D., B. HOLMSTROM, AND P. MILGROM (1990): "Short–Term Contracts and Long–Term Agency Relationships," *Journal of Economic Theory*, 51(1), 1–31.

GRAY, J. A. (1978): "On Indexation and Contract Length," *Journal of Political Economy*, 86(1), 1–18.

GROSSMAN, S., AND O. HART (1983): "An Analysis of the Principal–Agent Problem," *Econometrica*, 51(1), 7–45.

HARRIS, M., AND B. HOLMSTROM (1987): "On the Duration of Agreements," *International Economic Review*, 28(2), 389–406.

HOLMSTROM, B. (1979): "Moral Hazard and Observability," *The Bell Journal of Economics*, 10, 74–91.

KREPS, D. M. (1990): "Corporate Culture and Economic Theory," in *Perspectives on Positive Political Economy*, ed. by J. Alt, and K. Shepsle, pp. 90–143. Cambridge University Press.

KREPS, D. M., AND R. WILSON (1982): "Reputation and Imperfect Information," *Journal of Economic Theory*, 27(2), 253–279.

LAVER, M., AND W. B. HUNT (1992): *Policy and Party Competition.* Routledge, New York and London.

LIN, X. (1999): "Central–Bank Independence, Economic Behavior, and Optimal Term Lengths: Comment," *The American Economic Review*, 89(4), 1056–1062.

LINDNER, A. (2000): “Long–term appointment of central bankers: costs and benefits,” *European Journal of Political Economy*, 16, 639–654.

MacLEOD, B. W., and J. M. MALCOMSON (1989): “Implicit Contracts, Incentive Compatibility, and Involuntary Unemployment,” *Econometrica*, 57(2), 447–480.

MALCOMSON, J. D., and F. SPINNEWYN (1988): “The Multiperiod Principal–Agent Problem,” *Review of Economic Studies*, 55, 391–408.

MILGROM, P. R., and J. ROBERTS (1982): “Predation, Reputation, and Entry Deterrence,” *Journal of Economic Theory*, 27(2), 280–312.

MIRRLEES, J. A. (1999): “The Theory of Moral Hazard and Unobservable Behaviour: Part I,” *The Review of Economic Studies*, 66(226), 3–21.

MURPHY, K. J. (2000): “What effect does uncertainty have on the length of labor contracts?,” *Labour Economics*, 7, 181–201.

RADNER, R. (1981): “Monitoring Cooperative Agreements in a Repeated Principal–Agent Relationship,” *Econometrica*, 49(5), 1127–1148.

RICH, R., and J. TRACY (2004): “Uncertainty and Labor Contract Durations,” *The Review of Economics and Statistics*, 86(1), 270–287.

UNITED STATES DEPARTMENT OF LABOR (2006): “Employee Tenure in 2006,” *Labor Force Statistics from the Current Population Survey. News 06-1563.*

WALLACE, F. H. (2001): “The Effects of Shocks Size and Type on Labor–Contract Duration,” *Journal of Labour Economics*, 19(3), 658–681.

WALLER, C., and C. E. WALSH (1996): “Central–bank Independence, Economic Behavior and Optimal Term Lengths,” *American Economic Review*, 86(5), 1139–1153.

WOLDENDORP, J., H. KEMAN, and I. BUDGE (1998): “Party Government in 20 Democracies: an update (1990-1995),” *European Journal of Political Research*, 33, 125–164.

# Mathematical Appendices

## A The sequential equilibrium of the game

Proof of result 2.1.

The probability the third party attaches in period $k+1$ to the agent being trustworthy is $p_{k+1}$, which is computed according to Bayes' rule on the equilibrium strategy of the agent, assuming that a trustworthy agent always honors trust:

$$p_{k+1} = \begin{cases} p_k & \text{if } p_k > 0 \text{ and in period } k \text{ the agent is not given trust} \\ \frac{p_k}{p_k+(1-p_k)h_k} & \text{if } p_k > 0 \text{ and in period } k \text{ the agent honors trust} \\ 0 & \text{if } p_k = 0 \text{ or in period } k \text{ the agent abuses trust} \end{cases} \tag{A.1}$$

Let us now determine the unique sequential equilibrium of the game [38], starting from period $K$ and working backwards.

*Period $K$*

The expected return for the untrustworthy agent/responder $(R)$ in period $K$ if he is given the move is

$$J_R(K, p_K) = h_K + (1-h_K)\,o = o - (o-1)h_K$$

which is strictly decreasing in $h_K$. So, in the last period the agent abuses trust, *i.e.*, $h_K = 0$. The expected return for the third party/proposer $(P)$ is therefore

$$J_P(K, p_K) = t_K(p_K - (1-p_K)l) = t_K\,(p_K(1+l) - l)$$

The third party chooses $t_K$ that maximizes the above function. Hence, [39]

$$t_K = \begin{cases} 1 & \text{if } p_K > \frac{l}{1+l} \\ 0 & \text{if } p_K < \frac{l}{1+l} \end{cases} \tag{A.2}$$

[38] In fact, in the zero–probability event where $p$ and $l$ are such that $p = (\frac{l}{l+1})^k$ for some $k < K$, $k \in \mathbf{N}$, there exist multiple equilibria. We will henceforth rule out this case.

[39] When $p_K = \frac{l}{1+l}$, the third party is indifferent between its two pure strategies; we show below that, in this case, if $K > 1$, the unique sequential equilibrium calls for $t_K = \frac{o-1}{o}$.

*Period K-1.*

The continuation payoff in period $K$ for the untrustworthy agent is

$$V_R\left(K, p_K\right)=\begin{cases} o & \text{if } p_K>\frac{l}{1+l} \\ o\,t_K & \text{if } p_K=\frac{l}{1+l} \\ 0 & \text{if } p_K<\frac{l}{1+l}\end{cases} \tag{A.3}$$

The third party's continuation payoff is instead

$$V_P\left(K, p_K\right)=\begin{cases} p_K-l(1-p_K) & \text{if } p_K>\frac{l}{1+l} \\ 0 & \text{if } p_K\leq\frac{l}{1+l}\end{cases} \tag{A.4}$$

From EQS. (A.1) and (A.3), the agent's expected return in periods $K-1$ and $K$ *if* in period $K-1$ he is given the move is

$$J_R\left(K-1, p_{K-1}\right)=h_{K-1}(1+V_R(K, \frac{p_{K-1}}{p_{K-1}+(1-p_{K-1})h_{K-1}}))+(1-h_{K-1})\,o$$

which simplifies to

$$J_R\left(K-1, p_{K-1}\right)=o+h_{K-1}(V_R(K, \frac{p_{K-1}}{p_{K-1}+(1-p_{K-1})h_{K-1}})-(o-1)) \tag{A.5}$$

The following two results characterize the equilibrium in periods $K-1$ and $K$.

RESULT A.1 *If* $p_{K-1}>\frac{l}{1+l}$, *in equilibrium* $h_{K-1}=t_{K-1}=t_K=1$.

PROOF.
From EQ. (A.1), if in period $K-1$ the agent is given trust and he honors it, $p_K=\frac{p_{K-1}}{p_{K-1}+(1-p_{K-1})h_{K-1}}\geq p_{K-1}>\frac{l}{1+l}$. From EQ. (A.2), $t_K=1$; moreover, taking into account EQ. (A.3), from EQ. (A.5), $J_R\left(K-1, p_{K-1}\right)=o+h_{K-1}\left(o-(o-1)\right)=o+h_{K-1}$, which is strictly increasing in $h_{K-1}$; this proves that, in equilibrium, $h_{K-1}=1$. In period $K-1$, both types of agents honor trust, so the third party chooses to give trust, *i.e.*, $t_{K-1}=1$. ■

RESULT A.2 *If* $p_{K-1}\in(0,\frac{l}{1+l})$, *in equilibrium* $h_{K-1}=\frac{1-\frac{l}{l+1}}{\frac{l}{l+1}}\frac{p_{K-1}}{1-p_{K-1}}$; *if in period* $K-1$ *the agent honors trust, in equilibrium* $p_K=\frac{l}{1+l}$ *and* $t_K=\frac{o-1}{o}$.

PROOF.
If $h_{K-1} = 1$, from EQ. (A.1), $p_K = p_{K-1} < \frac{l}{1+l}$; hence, taking into account EQ. (A.3), from EQ. (A.5), $J_R(K-1, p_{K-1}) = o - h_{K-1}(o-1)$, which is strictly decreasing in $h_{K-1}$, so $h_{K-1} = 1$ cannot be optimal.
If $h_{K-1} = 0$, from EQ. (A.1), the third party attaches a probability $p_K = 1$ to a agent who honored trust being trustworthy; hence, taking into account EQ. (A.3), from EQ. (A.5), $J_R(K-1, p_{K-1}) = o + h_{K-1}(o - (o-1)) = o + h_{K-1}$, which is strictly increasing in $h_{K-1}$, so $h_{K-1} = 0$ cannot be optimal.
We have thus proved that, if the untrustworthy agent is given trust, $h_{K-1} \in (0,1)$, *i.e.*, in period $K-1$ he plays a mixed strategy. From EQ. (A.5), the untrustworthy agent does not choose $h_{K-1} = 1$ or $h_{K-1} = 0$ only when $V_R(K, \frac{p_{K-1}}{p_{K-1}+(1-p_{K-1})h_{K-1}}) = o - 1$, which, according to EQ. (A.3), occurs when $t_K = \frac{o-1}{o}$; but this requires that $p_K = \frac{l}{1+l}$; hence, from EQ. (A.1), $h_{K-1} = \frac{1-\frac{l}{l+1}}{\frac{l}{l+1}} \frac{p_{K-1}}{1-p_{K-1}}$. ∎

Taking into account RESULT A.2, the third party's expected return in period $K-1$ for the continuation of the game is [40]

$$J_P(K-1, p_{K-1}) = t_{K-1}(\frac{l+1}{l}p_{K-1}(1) + (1 - \frac{l+1}{l}p_{K-1})(-l)) + (1-t_{K-1})(0)$$

which simplifies to

$$J_P(K-1, p_{K-1}) = t_{K-1}(\frac{(l+1)^2}{l}p_{K-1} - l)$$

[40] In period $K-1$, the third party gives trust with probability $t_{K-1}$, and attaches a probability $p_{K-1} + (1-p_{K-1})t_{K-1} = \frac{l+1}{l}p_{K-1} > \mathbf{0}$ to the agent honoring trust, so it earns $\mathbf{1}$ in period $K-1$ and, from RESULT A.2 and EQ. (A.4), $\mathbf{0}$ in period $K$. And if the agent abuses trust, the third party earns $-l$ in period $K-1$ and, from EQS. (A.1) and (A.4), $\mathbf{0}$ in period $K$. Finally, with probability $(1-t_{K-1})$ in period $K-1$ the third party does not give trust, thus earning $\mathbf{0}$ both in period $K-1$ and, from EQS. (A.1) and (A.4), in period $K$.

The third party chooses $t_{K-1}$ that maximizes the above function. Hence, [41]

$$t_{K-1} = \begin{cases} 0 & \text{if } p_{K-1} < (\frac{l}{l+1})^2 \\ 1 & \text{if } p_{K-1} > (\frac{l}{l+1})^2 \end{cases}$$

The straightforward extension of the solution to earlier periods, together with EQ. (A.1), gives RESULT. 2.1. ∎

PROOF OF PROPOSITION 2.1

Let $\tilde{K}(p,l)$ be the integer part of $K(p,l)$, where $K(p,l) \in \mathbf{R}$ is the unique value that satisfies the equation $p = \left(\frac{l}{l+1}\right)^{K(p,l)-1}$. [42]

The proof of the proposition works in two steps.

*Step 1: Equilibrium in the trust game when $K < \tilde{K}(p,l)$*

RESULT A.3 *If $K < \tilde{K}(p,l)$, in equilibrium $t_k = 0$ for all $k = 1, 2, 3, \ldots, K$ and $h_k = \frac{p_k}{1-p_k} \frac{1-(\frac{l}{l+1})^{K-k}}{(\frac{l}{l+1})^{K-k}}$.*

PROOF.
The proof works by induction. Let $t_k = 0$, which calls for $p_k < (\frac{l}{l+1})^{K-k+1}$ (EQ. (2.1)). According to EQ. (2.3), $p_{k+1} = p_k$; hence, $p_{k+1} < (\frac{l}{l+1})^{K-k}$ and, from EQ. (2.1), $t_{k+1} = 0$. So, $t_k = 0 \Longrightarrow t_{k+s} = 0$ for all $s \in \mathbf{N}$. Now, we have to prove that $z_1 = 0$, which follows immediately from EQ. (2.1), since $K < \tilde{K}(p,l)$ implies that $p < \left(\frac{l}{l+1}\right)^K$. As for $h_k$, let $h_k = \frac{p_k}{1-p_k} \frac{1-(\frac{l}{l+1})^{K-k}}{(\frac{l}{l+1})^{K-k}}$, which calls for either $p_k \in (0, (\frac{l}{l+1})^{K-k})$ (EQ. (2.2)) or $p_k = 0$ (in which case, from EQ. (2.2), $h_k = 0$). According to EQ. (2.3),

[41] When $p_{K-1} = \mathbf{0}$, from EQ. (A.1), $p_K = \mathbf{0}$; hence, from EQS. (A.3) and (A.5), $J_R(K-1, \mathbf{0}) = o - (o-1)h_{K-1}$, which is strictly decreasing in $h_{K-1}$, so $h_{K-1} = \mathbf{0}$. From EQ. (A.4), $V_P(K, \mathbf{0}) = \mathbf{0}$. Hence, $J_P(K-1, \mathbf{0}) = t_{K-1}(-l+\mathbf{0}) + (1-t_{K-1})(\mathbf{0}+\mathbf{0})$, which is strictly decreasing in $t_{K-1}$, so $t_{K-1} = \mathbf{0}$. Finally, when $p_{K-1} = (\frac{l}{l+1})^2$, in period $K-1$ the third party is indifferent between giving and not giving trust; if $K > 2$, a slightly revised version of the proof of RESULT A.2 demonstrates that, in equilibrium, $t_{K-1} = \frac{o-1}{o}$.

[42] It is worth noting that $\tilde{K}(p,l) < K(p,l)$ because of the assumption $p = \left(\frac{l}{l+1}\right)^k$ for all $k < K$, $K \in \mathbf{N}$.

either $p_{k+1} = (\frac{l}{l+1})^{K-k}$ (so that $p_{k+1} < (\frac{l}{l+1})^{K-k-1}$) or $p_{k+1} = 0$; hence, from EQ. (2.2), $h_{k+1} = \frac{p_{k+1}}{1-p_{k+1}} \frac{1-(\frac{l}{l+1})^{K-k-1}}{(\frac{l}{l+1})^{K-k-1}}$ (which also holds for $p_{k+1} = 0$), so $h_k = \frac{p_k}{1-p_k} \frac{1-(\frac{l}{l+1})^{K-k}}{(\frac{l}{l+1})^{K-k}} \Longrightarrow h_{k+s} = \frac{p_{k+s}}{1-p_{k+s}} \frac{1-(\frac{l}{l+1})^{K-k-s}}{(\frac{l}{l+1})^{K-k-s}}$ for all $s \in \mathbf{N}$.

Finally, we have to prove that $y_1 = \frac{p}{1-p} \frac{1-(\frac{l}{l+1})^{K-1}}{(\frac{l}{l+1})^{K-1}}$, which follows immediately from EQ. (2.2), since $K < \tilde{K}(p,l)$ implies that $p < \left(\frac{l}{l+1}\right)^{K} < \left(\frac{l}{l+1}\right)^{K-1}$. ■

*Step 2: Equilibrium in the trust game when $K \geq \tilde{K}(p,l)$*

RESULT A.4 *If $K > \tilde{K}(p,l)$, $h_k = t_k = 1$ for all $k = 1, 2, \ldots, K - \tilde{K}(p,l)$*

PROOF.

From EQ. (2.3), as long as the agent is given the move and $y_i = 1$ for all $i \leq k$, we have $p_{k+1} = p$. From the definition of $\tilde{K}(p,l)$, $p \geq \left(\frac{l}{l+1}\right)^{K-k} \Longleftrightarrow k \leq K - \tilde{K}(p,l)$. The RESULT immediately follows from EQS. (2.1) and (2.2). ■

RESULT A.5 *If $K \geq \tilde{K}(p,l)$, in equilibrium $h_{K-\tilde{K}(p,l)+1} = \frac{p}{1-p} \frac{1-(\frac{l}{l+1})^{\tilde{K}(p,l)-1}}{(\frac{l}{l+1})^{\tilde{K}(p,l)-1}} < 1$ and $t_{K-\tilde{K}(p,l)+1} = 1$. Moreover, if in all previous periods trust was given and honored, $h_k = \frac{l}{l+1} \frac{1-(\frac{l}{l+1})^{K-k}}{1-(\frac{l}{l+1})^{K-k+1}} < 1$. Finally, $t_k = \frac{o-1}{o}$ for all $k = K - \tilde{K}(p,l) + 2, \ldots, K$ if in period $k-1$ the third party gave trust and the agent honored it, $t_k = 0$ otherwise.*

PROOF.

From the definition of $\tilde{K}(p,l)$ (with $\tilde{K}(p,l) < K(p,l)$), $p < \left(\frac{l}{l+1}\right)^{\tilde{K}(p,l)-1}$ and $p > (\frac{l}{l+1})^{\tilde{K}(p,l)}$. From EQ. (2.3) and RESULT A.4, $p_{K-\tilde{K}(p,l)+1} = p$. From EQ. (2.2), $p = p_{K-\tilde{K}(p,l)+1} < \left(\frac{l}{l+1}\right)^{\tilde{K}(p,l)-1} \Longrightarrow h_{K-\tilde{K}(p,l)+1} = \frac{p}{1-p} \frac{1-(\frac{l}{l+1})^{\tilde{K}(p,l)-1}}{(\frac{l}{l+1})^{\tilde{K}(p,l)-1}} < 1$; moreover, from EQ. (2.1), $p_{K-\tilde{K}(p,l)+1} > (\frac{l}{l+1})^{\tilde{K}(p,l)} \Longrightarrow t_{K-\tilde{K}(p,l)+1} = 1$.

The remaining part of the proof works by induction. Let $h_k = \frac{p_k}{1-p_k} \frac{1-(\frac{l}{l+1})^{K-k}}{(\frac{l}{l+1})^{K-k}}$.

From EQ. (2.3), if in period $k$ the third party gives trust and the agent honors it, then $p_{k+1} = (\frac{l}{l+1})^{K-k}$, which, taking into account to EQ. (2.1), implies that $t_{k+1} = \frac{o-1}{o}$. Moreover, $p_{k+1} < (\frac{l}{l+1})^{K-k-1}$; hence, from EQ. (2.2), $h_{k+1} = \frac{p_{k+1}}{1-p_{k+1}} \frac{1-(\frac{l}{l+1})^{K-k-1}}{(\frac{l}{l+1})^{K-k-1}} < 1$, which, taking into account that $p_{k+1} = (\frac{l}{l+1})^{K-k}$, simplifies to $h_{k+1} = \frac{l}{l+1} \frac{1-(\frac{l}{l+1})^{K-k-1}}{1-(\frac{l}{l+1})^{K-k}}$. Finally, if in period $k$ the third party does not give trust, from EQ. (2.3) we have $p_{k+1} = p_k < (\frac{l}{l+1})^{K-k}$. Hence, according to EQ. (2.1), $t_{k+1} = 0$. ■

PROP. 2.1 is thus proved. ■

PROOF OF RESULT 2.2

Let $\bar{k}$ be a period when the principal must hire a new agent. For a given $K$, the principal observes $\hat{\theta}_{\bar{k}}$ and chooses an agent $\theta^* = \theta(\hat{\theta}_{\bar{k}}, K)$. Given the assumptions on the distribution function of the desired characteristics, after a shock, *a priori* all agents perform equally well. As a consequence, for any $K$, $\theta^* = \hat{\theta}_{\bar{k}}$, *i.e.*, in period $\bar{k}$ the principal hires an agent who exactly matches her needs at that moment.
Let now arbitrarily choose an origin on the circle. The density function is $\frac{1}{2d}$. Once a shock has occurred, the per–period expected loss for the principal related to her agent's skills stays constant at

$$2 \int_{\hat{\theta}_{\bar{k}}}^{\hat{\theta}_{\bar{k}}+d} \frac{1}{2d} 2u(\hat{\theta}_k - \hat{\theta}_{\bar{k}}) d\hat{\theta}_k = 2\frac{1}{d} u \left[ \frac{(\hat{\theta}_k - \hat{\theta}_{\bar{k}})^2}{2} \right]_{\hat{\theta}_{\bar{k}}}^{\hat{\theta}_{\bar{k}}+d} = 2\frac{u}{d}\frac{d^2}{2} = ud$$

There is, however, a probability $\gamma^k$ that, at time $\bar{k} + k$, no shock has yet occurred, so that in that period the principal suffers no loss from continuing the relationship with the initial agent because the latter has exactly the characteristic she desires. As a consequence, the per–period expected utility at time $\bar{k}$ for the principal related to her agent's skills when the length of the delegation arrangement is $K$ periods is

$$\pi^F(K) = -ud + ud\frac{1}{K} \sum_{k=0}^{K-1} \gamma^k = -ud + \frac{1}{K} \frac{1-\gamma^K}{1-\gamma} ud$$

As for $\pi^T(K)$, from PROP. 2.1, in the first $K - \tilde{K}(p, l)$ periods trust is given and honored; hence, in these periods the principal earns the "cooperation" payoff 1. In the "transition" period, the third party gives trust and the agent starts playing a mixed strategy; this implies that in the last $\tilde{K}(p,l)$ periods the principal earns an overall payoff of $o$. Hence,

$$\pi^T(K) = \frac{(K - \tilde{K}(p,l)) + o}{K} = 1 - \frac{\tilde{K}(p,l) - o}{K}$$

As a consequence,

$$\pi(K) = \pi^T(K) + \pi^F(K) = 1 - \frac{\tilde{K}(p,l) - o}{K} - ud + \frac{1}{K}\frac{1-\gamma^K}{1-\gamma}ud$$

As for the properties of $\pi^F(K)$, from

$$\pi^{F\,\prime}(K) = \frac{1}{1-\gamma}ud(\frac{-K\gamma^K \ln\gamma - 1 + \gamma^K}{K^2})$$

$\pi^{F\,\prime}(K) < 0$ for all $K$ if (and only if) $K\ln\gamma - 1 > -\frac{1}{\gamma^K}$ for all $K$.
Let $\eta(K) = -\frac{1}{\gamma^K}$. The line which is tangent to $\eta(K)$ at $K = 0$ is $K\ln\gamma - 1$. Such line lies above the curve, since $\eta(K)$ is concave. This proves that $K\ln\gamma - 1 > -\frac{1}{\gamma^K}$ for all $K$. Hence, $\pi^{F\,\prime}(K) < 0$ for all $K$.
From

$$\pi^{F\,\prime\prime}(K) = \frac{1}{1-\gamma}ud(\frac{\gamma^K(-2 + 2K\ln\gamma - K^2(\ln\gamma)^2) + 2}{K^3})$$

$\pi^{F\,\prime\prime}(K) > 0$ for all $K$ if (and only if) $-2 + 2K\ln\gamma - K^2(\ln\gamma)^2 > -\frac{2}{\gamma^K}$ for all $K$.
From $\gamma^{-K} = 1 - K\ln\gamma + \frac{(K\ln\gamma)^2}{2!} - \frac{(K\ln\gamma)^3}{3!} + \ldots$,

$$-\frac{2}{\gamma^K} = -2 + 2K\ln\gamma - K^2(\ln\gamma)^2 + 2\frac{K^3(\ln\gamma)^3}{3!} - \ldots$$

From $2\frac{K^3(\ln\gamma)^3}{3!} < 0$, we have $-2 + 2K\ln\gamma - K^2(\ln\gamma)^2 > -\frac{2}{\gamma^K}$ for all $K$.
Hence, $\pi^{F\,\prime\prime}(K) > 0$ for all $K$.
The proof of the other properties of $\pi^F(K)$ and $\pi^T(K)$ is trivial. ■

Proof of proposition 2.2

Differentiating eq. (2.6),

$$\pi'(K) = \frac{1}{K^2}\frac{1}{1-\gamma}ud\left(\gamma^K(1-K\ln\gamma)-(1-(1-\gamma)\frac{\tilde{K}(p,l)-o}{ud})\right)$$

Letting $\beta = \dfrac{\tilde{K}(p,l)-o}{ud}$ and $g(K) = \gamma^K(1-K\ln\gamma)+(1-\gamma)\beta-1$, from $\dfrac{1}{K^2}\dfrac{1}{1-\gamma}ud > 0$ for all $K$, $sign(\pi'(K)) = sign(g(K))$. Hence, $\pi'(K) = 0 \Leftrightarrow g(K) = 0$.

Let now $\bar{K}$ be such that $g(\bar{K}) = 0$. From $g'(K) = -\gamma^K K(\ln\gamma)^2 < 0$, whenever $\bar{K}$ exists, it is unique.

From $\lim_{K\to\infty} g(K) = (1-\gamma)\beta - 1$, if $\gamma \leq 1-\dfrac{1}{\beta}$, $\lim_{K\to\infty} g(K) \geq 0$. Taking into account that $g'(K) < 0$ for all $K$, $\lim_{K\to\infty} g(K) \geq 0 \Rightarrow g(K) > 0$ for all $K$. Hence, $\bar{K}$ does not exist and, for $K \geq \tilde{K}(p,l)$, $\pi(K)$ is strictly increasing in $K$. As a consequence, the optimal length of the delegation arrangement for the principal $K^*$ is either infinite ($K^R = \infty$) or one period ($K^N = 1$). This calls to compare $\pi(\infty) = \lim_{K\to\infty}\pi(K) = 1 - ud$ and $\pi(1) = 0$.

Finally, if $\gamma > 1-\dfrac{1}{\beta}$, $\lim_{K\to\infty} g(K) < 0$. From $g(0) = (1-\gamma)\beta > 0$ and the intermediate value theorem, $\bar{K}$ always exists. From $g'(K) < 0$ for all $K$, $g(K) > 0$ for all $K < \bar{K}$ and $g(K) < 0$ for all $K > \bar{K}$. Hence, for $K \geq \tilde{K}(p,l)$, $\pi(K)$ reaches its maximum at $K^R = \bar{K}$ if $\bar{K} \geq \tilde{K}(p,l)$ and at $K^R = \tilde{K}(p,l)$ if $\bar{K} < \tilde{K}(p,l)$ — more concisely, for $K \geq \tilde{K}(p,l)$, $\pi(K)$ reaches its maximum at $K^R = \max\{\tilde{K}(p,l),\bar{K}\}$. As a consequence, the optimal length $K^*$ of the delegation arrangement for the principal is either $K^R = \max\{\tilde{K}(p,l),\bar{K}\}$ or $K^N = 1$. This calls to compare $\pi(\max\{\tilde{K}(p,l),\bar{K}\})$ and $\pi(1) = 0$. ■

# B Comparative statics

## B.1 Costs and benefits of a stable relationship

PROOF OF PROPOSITION 3.1

The proof of the proposition works in three steps..

*Step 1: Existence of $\bar{K}$ and comparative statics.*

RESULT B.1 *$\bar{K}$ exists if (and only if) $\beta < \frac{1}{1-\gamma}$, it tends to zero as $\beta$ tends to zero, it is strictly increasing in $\beta$ and tends to infinite as $\beta$ tends to $\frac{1}{1-\gamma}$.*

PROOF

For a given $\gamma$, PROP. 2.2 also determines the domain of $\bar{K} \geq 0$ as a function of $\beta : \mathfrak{F} = \{\beta \,|\beta > 0 \text{ and } \beta < \frac{1}{1-\gamma}\}$. For any $\gamma \in (0,1)$, $\mathfrak{F} = \emptyset$.
Now, we study the behavior of $\bar{K}$ at the boundaries of its domain. According to PROP. 2.2, $\bar{K}$ is the solution to

$$\gamma^K (1 - K \ln \gamma) = 1 - \beta + \gamma\beta \tag{A.6}$$

It is easy to see that the LHS of EQ. (A.6) is decreasing in $K$, tends to one as $K$ tends to zero and tends to zero as $K$ tends to infinite.
For any $\gamma \in (0,1)$, $\inf \mathfrak{F} = 0$. As $\beta$ tends to zero, the RHS of EQ. (A.6) tends to one; in order for also the LHS to tend to one, $K$ must tend to zero. Hence, $\bar{K}$ tends to zero as $\beta$ tends to zero.
For a given $\gamma$, $\sup \mathfrak{F} = \frac{1}{1-\gamma}$. As $\beta$ tends to $\frac{1}{1-\gamma}$, the RHS of EQ. (A.6) tends to zero; in order for also the LHS to tend to zero, $K$ must tend to infinite. Hence, $\bar{K}$ tends to infinite as $\beta$ tends to $\frac{1}{1-\gamma}$.
Finally, let us consider a pair $(\beta, \bar{K})$ that satisfies EQ. (A.6). By totally differentiating it, $\frac{d\bar{K}}{d\beta} = \frac{1-\gamma}{K\gamma^K (\ln \gamma)^2} > 0$. ■
It follows that $\bar{K}$ is strictly decreasing in $ud$.

*Step 2: Comparative statics on $K^R$.*

When $K \geq \tilde{K}(p,l)$, from PROP. 2.2 $\pi(K)$ reaches its maximum at

$$K^R = \begin{cases} \max\{\tilde{K}(p,l), \bar{K}\} & \text{if } \beta < \frac{1}{1-\gamma} \\ \infty & \text{if } \beta \geq \frac{1}{1-\gamma} \end{cases} \tag{A.7}$$

$\bar{K}$ is strictly increasing in $\beta$, so it is strictly decreasing in $ud$ (Step 1). On the other hand, $\tilde{K}(p,l)$ does not depend on $ud$. Hence, from EQ. (A.7), it follows immediately that $K^R$ is non–increasing (henceforth decreasing) in $ud$.

*Step 3: Comparative statics on $K^*$.*

Let us consider the two–stage procedure for finding $K^*$. Let us initially focus on the case where in the first stage the principal chose to let her agent build a reputation for trustworthiness by setting $K \geq \tilde{K}(p,l)$. As $ud$ tends to zero (and, hence, $\beta$ tends to infinite), $\bar{K}$ does not exist. Thus, $K^R = \infty$, and $\pi(K^R)$ tends to one. Both $\pi(\infty) = 1 - ud$ and $\pi(\max\{\bar{K}, \tilde{K}(p,l)\})$ (when $\bar{K}$ exists) are strictly decreasing in $ud$ and tend to $-\infty$ as $ud$ tends to infinite. Hence, $\pi(K^R)$ tends to one as $ud$ tends to zero, it is strictly decreasing in $ud$ and tends to $-\infty$ as $ud$ tends to infinite. Let us consider now the case where the principal renounced reputation–building and set $K < \tilde{K}(p,l)$. In this case, $\pi(K) = \pi^F(K)$ reaches its maximum at $K^N = 1$; the principal earns $\pi(1) = 0$, whatever the value of $ud$.
The principal solves the first stage by comparing $\pi(K^R)$ and $\pi(K^N) = \pi(1) = 0$. The proposition follows immediately from Step 2 ($K^R$ is decreasing in $ud$) and from the monotonic relationship between $\pi(K^R)$ and $ud$, taking into account that $\pi(K^R) > \pi(1) = 0$ (and, hence, $K^* = K^R$) as $ud$ tends to zero and $\pi(K^R) < \pi(1) = 0$ (and, hence, $K^* = K^N = 1$) as $ud$ tends to infinite. ■

PROOF OF PROPOSITION 3.2

From RESULT B.1, $\bar{K}$ is strictly increasing in $\beta$, so it is also strictly increasing in $\tilde{K}(p,l) - o$. As for $K^R$, from EQ. (A.7), taking into account that $\tilde{K}(p,l)$ is non–decreasing (henceforth increasing) in $\tilde{K}(p,l) - o$ [43], we have that $K^R$ is increasing in $\tilde{K}(p,l) - o$.
Let us now consider the two–stage procedure for finding $K^*$. If $ud < 1$, $\pi(\infty) > \pi(1)$ for all $\tilde{K}(p,l) - o$. Hence, $K^* = K^R \geq \tilde{K}(p,l)$ for all $\tilde{K}(p,l) - o$. The monotonic relationship between $K^*$ and $\tilde{K}(p,l) - o$ follows immediately from the monotonic relationship between $K^R$ and $\tilde{K}(p,l) - o$. On the other hand, if $ud \geq 1$, $\pi(\infty) \leq \pi(1)$. As $\tilde{K}(p,l) - o$ tends to infinite (and, hence,

[43] When we consider an increase in $\tilde{K}(p,l) - o$, we allow for either an increase in $\tilde{K}(p,l)$ for a given $o$ or a decrease in $o$ for a given $\tilde{K}(p,l)$, whereas we do not allow for a simultaneous change in both $\tilde{K}(p,l)$ and $o$.

$\beta$ tends to infinite), $\bar{K}$ does not exist, so $K^R = \infty$ and $\pi(K^R) \leq \pi(K^N) = \pi(1) = 0$. As $\tilde{K}(p,l) - o$ tends to zero [44] (and, hence, $\beta$ tends to zero), $\bar{K}$ tends to zero and, hence, $K^R = \max\{\bar{K}, \tilde{K}(p,l)\} = \tilde{K}(p,l)$. Thus, as $\tilde{K}(p,l) - o$ tends to zero, $K^* = K^N = 1$ if $\pi(\tilde{K}(p,l)) < 0$ and $K^* = K^R = \tilde{K}(p,l)$ if $\pi(\tilde{K}(p,l)) \geq 0$.
$\pi(K^R)$ is decreasing in $\tilde{K}(p,l) - o$ and $\pi(K^R) < \pi(K^N) = \pi(1)$ as $\tilde{K}(p,l) - o$ tends to infinite [45]. Hence, if $\pi(\tilde{K}(p,l)) < \pi(1)$, $K^* = K^N = 1$ for all $\tilde{K}(p,l) - o$ (we henceforth rule out this case), whereas if $\pi(\tilde{K}(p,l)) \geq \pi(1)$ there exists $\bar{\psi}$ such that $\pi(K^R) = \pi(1)$ when $\tilde{K}(p,l) - o = \bar{\psi}$, $\pi(K^R) > \pi(1)$ (and, hence, $K^* = K^R$) when $\tilde{K}(p,l) - o < \bar{\psi}$ and $\pi(K^R) < \pi(1)$ (and, hence, $K^* = K^N = 1$) when $\tilde{K}(p,l) - o > \bar{\psi}$. ■

## B.2 Stability of the environment

PROOF OF RESULT 3.1

For a given $\beta$, PROP. 2.2 determines the domain of $\bar{K} \geq 0$ as a function of $\gamma : \mathfrak{D} = \{\gamma \in (0,1) : \gamma > 1 - \frac{1}{\beta}\}$. For any $\beta > 0$, $\mathfrak{D} = \emptyset$.
We now study the behavior of the function $\bar{K}(\gamma)$ at the boundaries of its domain. When $\beta > 1$, $\inf \mathfrak{D} = 1 - \frac{1}{\beta}$. For a given $\gamma \in (0,1)$, the LHS of EQ. (A.6) tends to one as $K$ tends to zero, it is strictly decreasing in $K$ and tends to zero as $K$ tends to infinite. The right–side limit of the RHS of EQ. (A.6) when $\gamma$ tends to $1 - \frac{1}{\beta}$ is zero. With such a value of $\gamma$, the LHS of EQ. (A.6) also tends to zero only when $K$ tends to infinite. As a consequence, when $\beta > 1$, $\bar{K}(\gamma)$ tends to infinite as $\gamma$ tends to $1 - \frac{1}{\beta}$. On the other hand, when $\beta < 1$, $\gamma \in \mathfrak{D}$ for all $\gamma \in (0,1)$; hence, $\bar{K}(\gamma)$ always exists and $\inf \mathfrak{D} = 0$. The right–side limit of $\pi(K)$ as $\gamma$ tends to zero is $1 - ud + \frac{1}{K}(ud - (\tilde{K}(p,l) - o))$, which is strictly decreasing in $K$ for all $\beta < 1$. As a consequence, when $\beta < 1$, $\bar{K}(\gamma)$ tends to zero as $\gamma$ tends to zero. Finally, for any $\beta$, $\sup \mathfrak{D} = 1$. The left–side limit of $\pi(K)$ as $\gamma$ tends to one is $1 - \frac{1}{K}(\tilde{K}(p,l) - o)$, which is strictly increasing in $K$, whatever the value

---

[44] In our setting, $\tilde{K}(p,l) - o$ tends to zero when either (for a given $o$) $\tilde{K}(p,l)$ tends to $o$ or (for a given $\tilde{K}(p,l)$) $o$ tends to $\tilde{K}(p,l)$.

[45] In the zero–probability case where $ud = 1$, for any finite $\tilde{K}(p,l) - o$ the principal can asymptotically earn $\pi(1)$ by setting $K = \infty$. We avoid discussing whether this may also be the case when $\tilde{K}(p,l) - o$ tends to infinite, and simply assume that when $ud = 1$ the principal prefers $K = 1$ to $K = \infty$.

of $\beta$. As a consequence, for all $\beta > 0$, $\bar{K}(\gamma)$ tends to infinite as $\gamma$ tends to one. ■

PROOF OF PROPOSITION 3.3

When $K \geq \tilde{K}(p,l)$, from PROP. 2.2, $\pi(K)$ reaches its maximum at

$$K^R = \begin{cases} \infty & \text{if } \gamma \leq 1 - \frac{1}{\beta} \\ \max\{\tilde{K}(p,l), \bar{K}(\gamma)\} & \text{if } \gamma > 1 - \frac{1}{\beta} \end{cases} \tag{A.8}$$

From RESULT 3.2, when $\beta \geq 1$, $\bar{K}(\gamma)$ is not monotonic in $\gamma$. Moreover, $\tilde{K}(p,l)$ does not depend on $\gamma$. Hence, from EQ. (A.8), $K^R$ is also not monotonic in $\gamma$ (more precisely, decreasing in $\gamma$ for $\gamma < \bar{\gamma}$ and increasing in $\gamma$ for $\gamma > \bar{\gamma}$). Let us consider now the two–stage procedure for finding $K^*$. If $ud < 1$, $\pi(\infty) > \pi(1)$ for all $\gamma \in (0,1)$. It follows that $K^* = K^R$ for all $\gamma$, so the non–monotonic relationship between $K^*$ and $\gamma$ follows immediately from the non–monotonic relationship between $K^R$ and $\gamma$. On the other hand, if $ud \geq 1$, $\pi(\infty) \leq \pi(1)$. For all $\gamma \in (0, 1-\frac{1}{\beta}]$, $K^R = \infty$, so $K^* = K^N = 1$. Let us consider now $\gamma \in (1-\frac{1}{\beta}, 1)$, so $K^R = \max\{\tilde{K}(p,l), \bar{K}(\gamma)\}$. According to RESULT 3.1, as $\gamma$ tends to $1-\frac{1}{\beta}$, $\bar{K}(\gamma)$ tends to $\infty$ (and, hence, $K^R$ also tends to infinite). Hence, as $\gamma$ tends to $1-\frac{1}{\beta}$, $\pi(K^R)$ tends to $1 - ud \leq \pi(K^N) = \pi(1) = 0$. Moreover, as $\gamma$ tends to one, $\pi(K)$ tends to $1 - \frac{1}{K}(\tilde{K}(p,l) - o)$, which, taking into account that $K^R = \bar{K}(\gamma)$ tends to infinite as $\gamma$ tends to one, tends to $1 > \pi(K^N) = \pi(1) = 0$. Hence, from (i) $\pi(K^R) \leq \pi(K^N) = \pi(1) = 0$ as $\gamma$ tends to $1 - \frac{1}{\beta}$, (ii) $\pi(K^R)$ strictly increasing in $\gamma$ and (iii) $\pi(K^R) > \pi(K^N) = \pi(1) = 0$ as $\gamma$ tends to 1, there exists $\gamma^R \in (1-\frac{1}{\beta}, 1)$ such that $\pi(K^R) = \pi(K^N) = \pi(1) = 0$ when $\gamma = \gamma^R$, $\pi(K^R) < \pi(K^N) = \pi(1) = 0$ (and, hence, $K^* = K^N = 1$) when $\gamma \in (1-\frac{1}{\beta}, \gamma^R)$ and $\pi(K^R) > \pi(K^N) = \pi(1) = 0$ (and, hence, $K^* = K^R = \max\{\tilde{K}(p,l), \bar{K}(\gamma)\}$) when $\gamma \in [\gamma^R, 1)$. As a conclusion, $K^* = K^N = 1$ for all $\gamma \in (0, \gamma^R)$ and $K^* = K^R = \max\{\tilde{K}(p,l), \bar{K}(\gamma)\}$ for all $\gamma \in [\gamma^R, 1)$. Finally, taking into account that $\tilde{K}\,(p,l)$ does not depend on $\gamma$, from RESULT 3.2, we have that if $\gamma^R \geq \bar{\gamma}$, $K^* = K^* = \max\{\tilde{K}(p,l), \bar{K}(\gamma)\}$ is increasing in $\gamma$ for all $\gamma \in [\gamma^R, 1)$, while if $\gamma^R < \bar{\gamma}$, $\max\{\tilde{K}(p,l), \bar{K}(\gamma)\}$ is decreasing in $\gamma$ for all $\gamma \in [\gamma^R, \bar{\gamma})$ and increasing in $\gamma$ for all $\gamma \in (\bar{\gamma}, 1)$. ■

Proof of proposition 3.4

From result 3.3, when $\beta < 1$, $\bar{K}(\gamma)$ is increasing in $\gamma$. Moreover, $\tilde{K}(p,l)$ does not depend on $\gamma$. Hence, from eq. (A.8), $K^R$ is also increasing in $\gamma$. Let us consider now the two–stage procedure for finding $K^*$. If $ud < 1$, $\pi(\infty) > \pi(K^N) = \pi(1) = 0$ for all $\gamma \in (0,1)$. It follows that $K^* = K^R$ for all $\gamma$, The monotonic relationship between $K^*$ and $\gamma$ follows immediately from the monotonic relationship between $K^R$ and $\gamma$. On the other hand, if $ud \geq 1$, as $\gamma$ tends to zero, $\bar{K}(\gamma)$ tends to zero; hence, $K^R = \max\{\tilde{K}(p,l), \bar{K}(\gamma)\}$ tends to $\tilde{K}(p,l)$, and $\pi(\tilde{K}(p,l))$ tends to $1 - ud + \frac{1}{\tilde{K}(p,l)}(ud - (\tilde{K}(p,l) - o)) < \pi(K^N) = \pi(1) = 0$ (when $\beta < 1$). Moreover, as $\gamma$ tends to one, $\pi(K)$ tends to $1 - \frac{1}{K}(\tilde{K}(p,l) - o)$, which, taking into account that $K^R = \bar{K}(\gamma)$ tends to infinite as $\gamma$ tends to one, tends to $1 > \pi(K^N) = \pi(1) = 0$. Hence, from (i) $\pi(K^R) \leq \pi(K^N) = \pi(1)$ as $\gamma$ tends to 0, (ii) $\pi(K^R)$ strictly increasing in $\gamma$ and (iii) $\pi(K^R) > \pi(K^N) = \pi(1) = 0$ as $\gamma$ tends to 1, there exists $\gamma^R \in (0,1)$ such that $\pi(K^R) = \pi(K^N) = \pi(1) = 0$ when $\gamma = \gamma^R$, $\pi(K^R) < \pi(K^N) = \pi(1)$ (and, hence, $K^* = K^N = 1$) when $\gamma \in (0, \gamma^R)$ and $\pi(K^R) > \pi(K^N) = \pi(1) = 0$ (and, hence, $K^* = K^R = \max\{\tilde{K}(p,l), \bar{K}(\gamma)\}$) when $\gamma \in [\gamma^R, 1)$. As a conclusion, $K^* = K^N = 1$ for all $\gamma \in (0, \gamma^R)$ and $K^* = K^R = \max\{\tilde{K}(p,l), \bar{K}(\gamma)\}$ increasing in $\gamma$ for all $\gamma \in [\gamma^R, 1)$ ■

## Quaderni dell'Istituto di economia internazionale, delle istituzioni e dello sviluppo dell'Università Cattolica del Sacro Cuore

(DAL 2002 QUADERNI DEL DIPARTIMENTO)

| | |
|---|---|
| 9401 | Beretta C. *"Is economic theory up to the needs of ethics?"* (Part I) (trad. it. "Le scelte individuali nella teoria economica" pubblicata in M. Magrin (a cura di) (1996) "La coda di Minosse", Franco Angeli, Milano) |
| 9402 | Beretta C. *"Alcune radici del problema dell'autonomia individuale"* |
| 9403 | Beretta C. *"Asimmetrie informative ed autonomia: le strutture contrattuali e la formazione dei mercati"* (Parte I) |
| 9404 | Merzoni G. *"Delega strategica e credibilità delle minacce nella contrattazione tra sindacato e impresa"* |
| 9405 | Beretta C. *"Alcune funzioni e caratteristiche delle regole"* (pubblicato in Rivista Internazionale di Scienze Sociali, a. CII, n. 3, luglio-settembre, pagg. 339-55) |
| 9501 | Beretta C. *"Having alternatives, being free and being responsible"* (pubblicato in Cozzi T. - Nicola P.C. - Pasinetti L.L. - Quadrio Curzio A. (a cura di) "Benessere, equilibrio e sviluppo. Saggi in onore di Siro Lombardini", Vita e Pensiero, Milano) |
| 9502 | Beretta C. - Beretta S. *"Il mercato nella teoria economica"* (pubblicato in Persone & Imprese, n. 2, 1995) |
| 9503 | Beretta S. - Fortis M. - Draetta U. *"Economic Regionalism and Globalism"* (Europe-Iran Roundtable, Third Session, may 26, 1995) |
| 9504 | Beretta S. *"World Trade Organization: Italia ed Europa nel nuovo assetto globale"* (pubblicato su Rivista Internazionale di Scienze Sociali, a. CIII, n. 3, luglio-settembre 1995, p. 415-456) |
| 9505 | Colangelo G. - Galmarini U. *"Ad Valorem Taxation and Intermediate Goods in Oligopoly"* |
| 9601 | Beretta S. *"Disavanzi correnti e movimenti finanziari. Una survey molto selettiva e qualche (ragionevole) dubbio"* |
| 9602 | Beretta C. *"Strumenti per l'analisi economica - 1"* |
| 9603 | Beretta C. *"Dottrina sociale della Chiesa e teoria economica"* |
| 9604 | Venturini L., *"Endogenous sunk costs and structural changes in the Italian food industry"* |

| | |
|---|---|
| 9701 | Natale P., *"Posted Vs. Negotiated Prices under Incomplete Informaion"* |
| 9702 | Venturini L. - Boccaletti S. - Galizzi G., *"Vertical Relationships and Dual Branding Strategies in the Italian Food Industry"* |
| 9703 | Pieri R., Rama D., Venturini L., *"Intra-Industry Trade in the European Dairy Industry"* |
| 9704 | Beretta C., *"Equilibrio economico generale e teoria dei contratti"* (pubblicato in Istituto Lombardo - Accademia di Scienze e Lettere, Incontro di studio n. 14, Disequilibrio ed equilibrio economico generale, Milano, 1998) |
| 9705 | Merzoni G., *"Returns to Process Innovation and Industry Evolution"* |
| 9801 | Beretta C., Beretta S., *"Footpaths in trade theory: Standard tools of analysis and results from general equilibrium theory"* |
| 9802 | Beretta C., *"Alcuni problemi di giustizia, dal punto di vista dell'economista"* |
| 9803 | Beretta C., *"La scelta in economia"* |
| 9901 | Merzoni G., *"Observability and Co-operation in Delegation Games: the case of Cournot Oligopoly"* |
| 9902 | Beretta C., *"Note sul mercantilismo e i suoi antecedenti"* |
| 9903 | Beretta C., *"A Ricardian model with a market for land"* |
| 0001 | Beretta S., *"Disavanzi nei pagamenti e commercio intertemporale: alcuni spunti di analisi 'reale'"* |
| 0002 | Beretta S., *"Strumenti finanziari derivati, movimenti di capitale e crisi valutarie degli anni Novanta: alcuni elementi per farsi un'idea"* |
| 0003 | Merzoni G., *"Stategic Delegation in Firms and the Trade Union"* |
| 0101 | Colombo F. – Merzoni G., *"Reputation, flexibility and the optimal length of contracts"* |
| 0102 | Beretta C., *Generalità sulla scelta in condizioni di certezza* |
| 0103 | Beretta C., *"L'ipotesi di completezza e le sue implicazioni"* |
| 0104 | Beretta C., *"Una digressione sulle implicazioni della completezza"* |
| 0201 | Beretta C., *"L'ipotesi di transitività"* |
| 0202 | Beretta C., *"Un'introduzione al problema delle scelte collettive"* |
| 0203 | Beretta C., *"La funzione di scelta"* |
| 0204 | Beretta C., *"Cenni sull'esistenza di funzioni indice di utilità"* |

0205 Colombo F. – Merzoni G., *"In praise of rigidity: the bright side of long-term contracts in repeated trust games"*
0206 Quadrio Curzio A., *"Europa: Crescita, Costruzione e Costituzione"*

## Quaderni editi da Vita e Pensiero*

0401 Uberti T. E., *"Flussi internazionali di beni e di informazioni: un modello gravitazionale allargato"*
0402 Uberti T. E. e Maggioni M. A., *"Infrastrutture ICT e relazionalità potenziale. Un esercizio di "hyperlinks counting" a livello sub-nazionale"*
0403 Beretta C., *"Specializzazione, equilibrio economico ed equilibrio politico in età pre-moderna"*
0404 Beretta C, *"L'esperienza delle economie 'nazionali'"*
0405 Beretta C. e Beretta S. , *"L'ingresso della Turchia nell'Unione Europea: i problemi dell'integrazione fra economie a diversi livelli di sviluppo"*
0406 Beretta C. e Beretta S., *"L'economia di Robinson"*
0501 Beretta C., *"Elementi per l'analisi di un sistema economico"*
0502 Beretta C., *"Mercato, società e stato in un'economia aperta - Parte I"*
0503 Beretta C., *"Mercato, società e stato in un'economia aperta - Parte II"*
0601 Beretta C., *"L'ipotesi di razionalità: Parte I"(*)*
0602 Beretta C., *"L'ipotesi di razionalità: Parte II"(*)*
0603 Beretta C., *"Can Common knowledge of rationality make information incomplete? The case of the centipede"(*)*
0604 Beretta C., *"Can Common knowledge of rationality make information incomplete? The case of the finitely repeated presoners' dilemma"(*)*

---

* Nuova linea di Quaderni DISEIS stampata grazie ad un accordo con l'Editrice Vita e Pensiero dell'Università Cattolica.

(*) Testo disponibile presso il DISEIS

Finito di stampare
nel mese di dicembre 2007
da Gi&Gi srl - Triuggio (MI)

ISBN 978-88-343-1647-4