Moral Hazard Model with Short-run Impatience and Imperfect Self-awareness *

Jianye YAN†

*I thank Prof. Jean TIROLE for his early advices on this topic. And I am deeply indebted to my advisor, Prof. Thomas MARIOTTI, for his indispensable instructions and discussions with me on the paper. I am also grateful to helpful comments from participants in the 25th Annual Congress of EEA and the 10th World Congress of the Econometric Society. The conversation with Sanxi LI is also acknowledged. Of course all the possible faults and defects inside the paper are my own.

†Toulouse School of Economics. Address: MH 208, TSE, Manufacture de Tabacs, 21 allées de Brienne, 31000, Toulouse, France. Phone: 0033-6-62608675. Email: yanjianye@gmail.com.

Abstract

We model a contract design problem in which a risk-neutral and time-consistent principal owns a risky project that requires work effort by a risk-averse agent who suffers from self-control problem and exhibits short-run impatient time preference. The effort of the agent is unobservable to the principal so the possibility of moral hazard arises. We investigate the willingness of the principal to arrange a long-run employment contract instead of a short-run one in order to circumvent the short-run impatience and take advantage of the long-run patience of the agent. If the principal has exogenous commitment power toward the agent, then the fact of advancing recruitment happens. And in this case, some traditional consensus on moral hazard model is violated, where the underlying insight is that the extent of present bias can be regarded as a complement to the risk tolerance, that is to say, when the agent is not too risk-averse, the optimal arrangement for the risk-neutral principal is to sell the risky asset to the risk-averse and short-run impatient agent, and just get a fixed dividend herself as a seemingly fully insured risk-averse investor. And when the agent involves some extent of naivety, the complementarity between time-inconsistency and risk tolerance is strengthened. On the behalf of the principal, the presence of the agents naivety is always good news but its impact on the social welfare is ambiguous which depends on different ways of defining consumer surplus.

Key words: *Moral hazard*; *Time-inconsistency*; *Short-run impatience*; *Naivety*; *Quasi-hyperbolic discounting*; *Risk attitude*; *Paternalism*

JEL classification: D03; D81; D82

1 Introduction

1.1 Time-inconsistency, Short-run Impatience, Self-awareness and Quasi-hyperbolic Discounting

In recent decades, economists began to pay attention to individuals' self control problem and model it inside pricing, contracting and mechanism design environments to see whether there exists radical impact to some traditional economic outcomes due to this new introduction of psychologically irrational aspect of agents. The terminologies such as time-inconsistency, short-run impatience, present-biased preferences and so on well up in the literature. The inconsistency of the agent's attitude or inter-temporal preferences toward future between today and tomorrow depicts the basic feature of self-control problem; in other word, in the long run, people are normally more patient or more rational to resist temptation, while in the short run, they are likely to be more impatient or may result in some decision inclined to some allurement. In the literature, there are mainly two streams of approaches to model self control problem. One is Gul-Pesendorfer temptation preference formulation (Gul-Pesendorfer, 2004). Another line is the quasi-hyperbolic discounting which originates from Strotz (1956), experiences the experimental conduction by Chung-Herrnstein (1961) and becomes generally and classically used after Phelps-Pollak (1968) and Laibson (1997). Sometimes some authors combine the two approaches to show the more compatibility of their models, like Eliaz-Spiegler (2006), who used a pair of function (u, v) representing the preferences of today and tomorrow.

In this paper, we employ quasi-hyperbolic discounting approach to express time-inconsistency/short-run impatience since it is more appropriate to our parameterization environment. The main idea of quasi-hyperbolic discounting is an introduction of a single parameter β inside the traditional exponential δ discounting; this β represents all the indispensable feature of present-biased time-inconsistency; so the (β, δ) preference may also be short for quasi-hyperbolic discounting in the subsequence. In this environment, an agent at time t treats the future (the present value of future utilities) as $u_t = v_t + \beta \sum_{s=t+1}^{\infty} \delta^{s-t} v_s$, where v represents the instantaneous period utility. $\beta < 1$ expresses the fact of short-run impulsion; if $\beta = 1$, the formulation degenerates to the traditional exponential time-consistent δ discounting, where the inter-temporal discounting trace (discount function) is exactly an exponential curve (solid line in Figure 1). When $\beta < 1$, the discounting tract at t approaches a hyperbola, which is the origin of the name "quasi-hyperbolic"; and its most important feature is the characterization of time-inconsistency caused by the short-run impatience, i.e., the different discounting trajectory between today and tomorrow (long and short dotted line in Figure 1, respectively).

Figure 1 is about here

Because of the difference between today and tomorrow, the self-image/self prediction toward tomorrow's extent of inconsistency becomes relevant. The terminologies like self-awareness, naivety or sophistication, constitute another dimension of concept in (β, δ) preference framework. For $\beta < 1$, we denote $\widehat{\beta}$ ($\beta \leq \widehat{\beta} \leq 1$) as the subjective prediction of her future short-run impulsive extent by the agent's present self. If $\widehat{\beta} = \beta$, we call the agent a sophisticate, that is, he has complete self-awareness of his self-control shortcoming. If $\widehat{\beta} = 1$, then this means that the agent is fully unaware of his future time-inconsistency, so we call it fully naive. For the intermediate case $\beta < \widehat{\beta} < 1$, it is called incomplete self-awareness or partial naivety in the literature (which is originated from O'Donoghue-Rabin (2001)), which means that the individual is somewhat over-confident toward his tomorrow's self-control ability.

1.2 Literature Review

After clearing out the terminologies in the time-inconsistent preference, we come to the question how organizations take into account the existence of self-control and naivety of the agents/employees/consumers and design optimal contracts on the behalf of the principal/employer/seller. After Laibson (1997) and Gul-Pesendorfer (2001), a bunch of papers arise. DellaVigna-Malmendier (2004) is typical of a literature which introduces quasi-hyperbolic discounting into traditional industrial organization and contract theory; they fit well the pricing behavior of health clubs with the present-biased preference of consumers and conclude that after taking into account the existence of time-inconsistency of individuals, the seemingly abnormal pricing scheme is found in fact optimal for the firm; but they did not set asymmetric information between the principal (firm) and the agent (consumer), that is, they are in the first-best world.

Following them, quite a few of authors investigate the second-best feature of contract design when agents bear self-control defect, like Esteban-Miyagawa (2006), Esteban-Miyagawa-Shum (2007) and Eliaz-Spiegler (2006), among which some are using Gul-Pesendorfer temptation preferences and some are belonging to quasi-hyperbolic discounting. But all of these papers are within adverse selection framework, i.e., the agent just holds some private information on her type and no hidden or unobservable behavior of the agent is involved.

To our knowledge, few of literature concern the moral hazard problem with a time-inconsistent agent. Gilpatric (2008) may be counted as a closer one since the author considers the shirking problem of quasi-hyperbolic agents.

But the model is not a standard moral hazard setting because there is no uncertainty between agent's effort and final output of the task, where the principal can verify perfectly the amount of effort exerted by the agent with postponement (ex post). Thus, Gilpatric (2008) should still be ranged inside the adverse selection branch of the literature since in that paper, the critical interest of the author is still that the principal is not sure about the time preference (type) of the agent (time-consistent or not, self-aware or not) ex ante.

1.3 Depiction and Main Findings of This Paper

Therefore, this paper aims at investigating the **moral hazard** contract between a time-consistent principal and a time-inconsistent agent. The employer owns a project that requires work effort by an agent today (Period 1) and will get an output tomorrow (Period 2). The effort level is not observable to the employer (the principal) and the output level is composed by the effort level plus a random shock. The private information feature of the effort level constitutes the moral hazard problem. And the compensation scheme to the agent can only be based on the output level. The employer is time-consistent[1] while the employee may be time-consistent or inconsistent,

[1]We can regard the employer as an organization and its corporate inter-temporal decision does not suffer from individual's self-control problem.

and even imperfectly self-aware when she is time-inconsistent, but the time preference parameters of the agent are public knowledge to the principal so the possibility of adverse selection is ruled out for the time being, i.e., we focus on pure moral hazard here.

The paper is organized as follows. Section 2 introduces the 2 players (risk-neutral principal and risk-averse agent) and gives 2 alternative contract outcomes in a 2-period model: (1) both principal and agent are time-consistent and patient enough, which is just a representation of traditional moral hazard problem; (2) the agent suddenly becomes short-run impatient but his long-run discounting factor keeps patient, i.e., time-inconsistency arises, then by encountering this, the employer's adjusted optimal contract and profit loss are shown. In Section 3, we raise the question whether the principal is interested in advancing the contract time, i.e., the employer starts recruitment from Period 0 but all the tasks take place in Period 1 (3-period long-run employment) instead of the former 2-period one, on the sake of taking advantage of long-run patience of the employee (δ) and circumventing her short-run impatience ($\beta\delta$). Section 4 involves the presence of agent's naivety and investigates its impact on the contracting scheme, the principal's profit and social welfare. Section 5 concludes.

We find that recruiting earlier brings strict amelioration to the principal if she has a commitment device toward the agent to prevent the resignation during the intermediate stage of the contract; but the profit loss due to the presence of employee's time-inconsistency cannot be fully alleviated, not like the conclusion in other literature dealing with adverse selection for a time-inconsistent agent, like Gilpatric (2008). A straightforward reason is that under the setting of moral hazard, the short-run private effort level cannot be stimulated by neither the principal nor the long-run agent's patient self.

In addition, in the optimal 3-period contract, some standard properties of moral hazard in Holmstrom (1979) and Harris-Raviv (1976) are violated, e.g., "no moral hazard problem under risk neutrality of the agent" and "slackness of non-manipulability constraints" no longer hold under the time-inconsistent agent environment. Combining these 2, we get an interesting finding: in this inter-temporal setting, the extent of short-run impatience can be regarded as a complement of risk tolerance. The upper binding non-manipulability constraint means that the principal sells the whole project to the agent and gets some fixed deposit from the agent, i.e., a risk-averse but time-inconsistent agent may be more appropriate to manage a risky asset than a risk-neutral but time-consistent person.

When the agent is imperfectly self-aware, the preceding finding is even strengthened. And the full naivety exhibits a perfect complementarity between the short-run impatience/time-inconsistency and the risk tolerance. At last, the concept of naivety in the quasi-hyperbolic discounting leaves a perplexing question: whether the absence of self-awareness is detrimental to the society or not? The answer depends on different kinds of definition of consumer surplus.

2 2-Period Benchmark Moral Hazard Contract

2.1 The Environment Setting and the Portrait of Principal-Agent

A principal has a project which results in tomorrow's single output y. The production function is $y = f(e, \varepsilon) = e + \varepsilon$. e is today's effort amount by an agent. And ε is a stochastic factor depending neither on the principal's nor on the agent's behavior; in addition, $E(\varepsilon) = 0$ and $Var(\varepsilon) = \sigma^2$. Here we take the simplest direct sum between e and ε to constitute the final output y for the sake of simplification of calculation in the whole paper. The key relationship between the principal and the agent is: the principal needs the agent's effort; the agent cannot realize the market value of y by himself so that he has to return it back to the principal and get some compensation. This constitutes the base of bilateral cooperation in the contract.

The timing of this 2-peirod model is depicted by Figure 2. In period 1, the principal and the agent sign the contract and in the same date the agent exerts his private effort e; tomorrow the principal gets her project's monetary output y and then compensates yesterday's effort of the agent in the amount of w_2. We call this timing as a benchmark since it is a short-run contract, by contrast later on we will extend to 3 periods.

Figure 2 is about here

For simplicity, in the following, we normalize the time-consistent (long-run) discounting factor δ to be 1[2]. The principal is set as time-consistent and risk-neutral, thus her preference at Period 1 can be expressed as $\pi = E(y - w_2)$. The cost (disutility) of the agent's effort takes the quadratic form: $C(e) = \frac{e^2}{2}$. As usual in moral hazard literature, the agent is set to be risk-averse; so in Period 2, the utility of the agent is $U_2 = E(w_2) - \frac{\gamma}{2} Var(w_2)$, in which γ measures the extent of agent's risk-aversion (the agent would be risk-averse unless $\gamma = 0$). We notice that the agent may involve a short-run impulsive parameter β $(0 < \beta \leq 1)$; therefore, inter-temporally, in Period 1, the agent has the preference

$$U_1 = \beta U_2 - C(e) = \beta \left(E(w_2) - \frac{\gamma}{2} Var(w_2) \right) - \frac{e^2}{2}$$

from where we can see if $\beta = 1$, the agent's time preference coincides with that of the principal.

[2]With a general $0 < \delta \leq 1$, no qualitative results in the paper will change.

Before going on, we firstly clear out the "first-best" and "second-best" concepts in this environment.

Definition 1 *first-best contracting environment*

Both e and y are publicly observed by the principal.

Definition 2 *second-best contracting environment*

e is the private information of the agent. The principal cannot observe the actual output generated from the risky project, neither. The evaluation and compensation to the agent can only base on an amount of "y" handed back by the agent to the principal.

2.2 Moral Hazard Contract for a Patient Agent

At first we look at the optimal recruitment contract design between a principal and an agent under equivalent and full patience ($\beta = 1$), which is also a short revision of standard properties in moral hazard problem.

2.2.1 First-best

When e and y are perfectly observed by the principal, she can directly assign a take-it-or-leave-it effort level inside the contract, thus the first-best contract will appear as a form $(w_2,\ e)$. The risk-neutral principal seeks an optimal contract from the following maximization.

$$\max_{e,w_2} \pi = E\left(y - w_2\right) = e - w_2 \tag{1}$$

s.t.

$$U_1 = w_2 - \frac{e^2}{2} \geq 0 \tag{2}$$

We see that in the first-best, w_2 is a fixed amount of payment corresponding to the exact assigned effort level, so the risk attitude γ of the agent does not enter (2). The optimal contract can be immediately got: $\left\{e_{[1]}^{FB} = 1, w_{2[1]}^{FB} = \frac{1}{2}\right\}$, where the superscript "FB" stands for "first-best", and the subscript "[1]" represents the fully patient time preference of the agent. And we can get the first-best profit of the principal and the joint principal-agent surplus (social welfare) as: $\pi_{[1]}^{FB} = S_{1[1]}^{FB} = \frac{1}{2}$, where $S_t = \pi + U_t$.

2.2.2 Second-best

If e and y cannot be observed by the principal (the entire production procedure is private information of the agent), then a moral hazard problem occurs; the principal can only design w_2 contingent on y (the handed in amount of project output). In this paper we restrict ourselves to linear incentive scheme[3]: $w_2(y) = a + by$. So the second-best contract can be ab-

[3]General incentive form $w_2\left(y\right)$ may be an extension after this paper; but here the simplest calculation involved by linear incentive sufficiently gives us the insights of impact

stracted as the form (a, b). In addition, the slope of the incentive should be restricted to $0 \leq b \leq 1$, namely, the "non-manipulability constraints". The reason is stated in the following lemma.

Lemma 1 ***Non-manipulability constraints and truth-reporting***

$0 \leq b \leq 1$ *is necessary and sufficient for an implementable and truth-reporting second-best contract.*

Proof of Lemma 1. Firstly, necessity. If $b < 0$ (negative marginal reward), the agent would always disguise the output level y toward 0; symmetrically, if $b > 1$, the agent would like to borrow from outside to inflate output level y toward $+\infty$ (there would be room for arbitrage).

Secondly, sufficiency. When $b > (=) \, 0$, there is no (strong) incentive to disguise the output; when $b \leq 1$, there is no room for profitable borrowing from outside. So facing an incentive contract with $0 \leq b \leq 1$, the truth-reporting strategy is (weakly) optimal for the agent. ■

Under $0 \leq b \leq 1$, the amount of e is still private information of the agent. And the fact that y contains an uncertain factor ε constitutes the possibility of shirking. The optimal moral hazard incentive scheme (a, b) can be solved by the backward induction in a dynamic game. First of all the

by the time-inconsistency of the agent on the form of moral hazard contract, the principal's project profit and the social welfare.

agent maximizes his utility by choosing e facing a and b:

$$\max_{e} U_1 = E\left(w_2\right) - \frac{\gamma}{2} Var\left(w_2\right) - C\left(e\right) = a + be - \frac{\gamma}{2} b^2 \sigma^2 - \frac{e^2}{2} \tag{3}$$

from which the best choice of the effort is $e = b$. Then the principal takes the agent's best reaction into account and seeks optimal a and b subject to the agent's self 1's participation and the non-manipulability constraints:

$$\max_{a,b} E\left(y - w_2\right) = \left(1 - b\right) E\left(y\right) - a = \left(1 - b\right) b - a \tag{4}$$

s.t.

$$U_1 = a + b^2 e - \frac{\gamma}{2} b^2 \sigma^2 - \frac{b^2}{2} \geq 0 \tag{5}$$

$$0 \leq b \leq 1 \tag{6}$$

from which the optimal second-best incentive and the associated economic outcomes are:

$$\begin{cases} b_{[1]}^{SB} = \frac{1}{1+\gamma\sigma^2} \\ e_{[1]}^{SB} = \frac{1}{1+\gamma\sigma^2} \\ \pi_{[1]}^{SB} = S_{1[1]}^{SB} = \frac{1}{2(1+\gamma\sigma^2)} \end{cases}$$

where the superscript "SB" stands for "second-best". Notice that if the agent is risk-neutral, i.e., $\gamma = 0$, we obtain: $e_{[1]}^{SB} = e_{[1]}^{FB}$, $\pi_{[1]}^{SB} = \pi_{[1]}^{FB}$, $S_{1[1]}^{SB} = S_{1[1]}^{FB}$, which is an important moral hazard problem property. And we review two standard results in moral hazard contract for patient agent established in Harris-Raviv (1976) and Holmstrom (1979) in the following proposition.

Proposition 1 ***Second-best moral hazard contract for a fully patient agent***

(i) The problem of moral hazard vanishes when the agent is risk-neutral: when $\gamma = 0$, all the second-best economic outcomes (including expected project output, the principal's profit and the joint principal-agent welfare) attain their first-best levels, respectively.

(ii) The non-manipulability constraints $0 \leq b \leq 1$ are not binding.

Proof of Proposition 1. It is sufficient to show the detailed solving procedure of (4) subject to (5)(6).

First we ignore the constraints $0 \leq b \leq 1$ and cancel a from binding the participation constraint, then we get $b = \frac{1}{1+\gamma\sigma^2}$, which belongs to $(0, 1]$, so the non-manipulability constraints are automatically satisfied (not binding): this verifies (ii).

By inserting $b = \frac{1}{1+\gamma\sigma^2}$ and combining the result of (3) ($e = b$), we get $e = \frac{1}{1+\gamma\sigma^2}$, $\pi = S_1 = \frac{1}{2(1+\gamma\sigma^2)}$. The coincidence between second-best and first-best under $\gamma = 0$ can be obtained by the comparison with the result of (1) subject to (2). This verifies (i). ■

2.3 Moral Hazard Contract for a Short-run Impatient Agent

2.3.1 First-best

When the short-run patience level of the agent becomes strictly less than 1, we just replace the participation constraint (2) by (7):

$$U_1 = \beta w_2 - \frac{e^2}{2} \geq 0 \tag{7}$$

To maximize (1) subject to (7) gives us $\left\{e = \beta, w_2 = \frac{\beta}{2}\right\}$, and by imposing this contract, we can get $\pi = \frac{\beta}{2}$: the principal's profit is eroded by the presence of short-run impatience of the agent. But this contract is not optimal on the behalf of the principal. If the principal is not financially constrained in advance, why not front-load the payment from w_2 to w_1 (in Period 1) since tomorrow's benefit is heavily discounted by the short-run impulsive agent so as to avoid the profit and efficiency loss in the contract? The answer is "yes", the optimal first-best program of the principal becomes:

$$\max_{e,w_1} \pi = E\left(y - w_1\right) = e - w_1 \tag{8}$$

s.t.

$$U_1 = w_1 - \frac{e^2}{2} \geq 0 \tag{9}$$

We see that the only difference between (8)(9) and (1)(2) is the replacement of w_2 by w_1, so the solution and all the economic outcomes coincide with the case of a fully patient agent: $e^{FB}_{[\beta]} = 1, \pi^{FB}_{[\beta]} = S^{FB}_{1[\beta]} = \frac{1}{2}$, where the subscript "$[\beta]$" stands for a $\beta < 1$ agent. So by rearranging the timing of the wage, the first-best allocation is not influenced by the fact whether the agent is more short-run impatient than the principal or not.[4]

2.3.2 Second-best

Before solving the optimal second-best contract scheme (a, b), we can expect that the detriment introduced by the short-run impatience of the agent to the principal and the economy; so similar logic raises: can we advance w_2 to w_1 to avoid this? Unfortunately, in the second-best world, the answer is "no". We state it more clearly in the following lemma and its proof.

Lemma 2 ***Infeasibility of*** w_1 ***in the second-best environment***

When e *is private information of the agent, one cannot front-load wage from Period 2 (*w_2*) to Period 1 (*w_1*).*

Proof of Lemma 2. y comes out one period later than the effort input, so w_1 has to be fixed and cannot be contingent on y (no incentive sense). If only facing a prepayment w_1, the complete shirk would happen, i.e., the private effort level $e = 0$ in Period 1, therefore the story ruins. ■

[4]The first-best profit's immunity from the agent's self-control defect is consistent with the findings in the classic literature regarding self-control/quasi-hyperbolic discounting like DellaVigna-Malmendier (2004).

So for a $\beta < 1$ agent, facing a contract scheme $(a,\ b)$, his self 1's utility maximization changes from (3) to (10):

$$\max_{e} U_1 = \beta\left(E\left(w_2\right) - \frac{\gamma}{2}Var\left(w_2\right)\right) - C\left(e\right) = \beta\left(a + be - \frac{\gamma}{2}b^2\sigma^2\right) - \frac{e^2}{2} \quad (10)$$

from which $e = \beta b$. By backward induction, the principal's problem can be stated as maximizing (11) subject to (12)(6), where (11) and (12) are, respectively:

$$\max_{a,b} E\left(y - w_2\right) = \left(1 - b\right)E\left(y\right) - a = \left(1 - b\right)\beta b - a \quad (11)$$

$$U_1 = \beta\left(a + b \cdot \beta b - \frac{\gamma}{2}b^2\sigma^2\right) - \frac{\left(\beta b\right)^2}{2} \geq 0 \quad (12)$$

The optimal b and the associated expected project output (effort level), the profit of the principal are:

$$\begin{cases} b^{SB}_{S[\beta]} = \frac{\beta}{\beta + \gamma\sigma^2} \\ e^{SB}_{S[\beta]} = \frac{\beta^2}{\beta + \gamma\sigma^2} \\ \pi^{SB}_{S[\beta]} = \frac{\beta^2}{2(\beta + \gamma\sigma^2)} \end{cases}$$

where the additional subscript "S" stands for "short-run". As for the consumer/agent surplus, since the agent is suffered from short-run impatience/self-control, we have two alternatives: one is the self subjective utility feeling of a short-run impatient agent; another one is a measure after the correction of short-run impatient time preference, which is from O'Donoghue-Rabin (2001). So we define the paternalism-type and let-it-be type consumer surplus as follows.

Definition 3 *(O'Donoghue-Rabin (2001))* ***Paternalism-type consumer surplus***

For a time-inconsistent/short-run impatient agent, his paternalism-type consumer surplus is the utility measure as if he were time-consistent/fully patient but suffered his time-inconsistent/short-run impatient decision: $\widetilde{U}_{t,P} = v_t + 1 \cdot \sum_{s=t+1}^{\infty} \delta^{s-t} v_s$, *where the superscript "˜" plus the subscript "P" represent "paternalism".*

We see from Definition 3 that the paternalism-type consumer surplus is an evaluation way after a paternalistic correction of the self-control problem.

Definition 4 ***Let-it-be type consumer surplus***

For a time-inconsistent/short-run impatient agent, the let-it-be-type consumer surplus is directly his own subjective utility measure.

Since the constraint (12) is binding, the let-it-be type consumer surplus is 0 and the let-it-be type social welfare (joint principal-agent surplus) is just composed of the principal's profit. While the paternalism-type consumer/agent surplus is the evaluation of the rationally patient society; under this concept, $\tilde{U}_{1,P}^{SB}(\beta) = a_{S[\beta]}^{SB} + b_{S[\beta]}^{SB} \cdot \beta b_{S[\beta]}^{SB} - \frac{\gamma}{2}\left(b_{S[\beta]}^{SB}\right)^2 \sigma^2 - \frac{\left(\beta b_{S[\beta]}^{SB}\right)^2}{2} = \frac{\beta^3(1-\beta)}{2(\beta+\gamma\sigma^2)^2} > 0 = U_1^{SB}(\beta)$.[5]

[5]This tells an interesting point: the agent may take advantage of the short-run impatience and preserve more surplus in the sense of "paternalism-type" evaluation under this

In the following proposition, we summarize the main features of the short-run second-best moral hazard contract.

Proposition 2 ***The impact of short-run impatience of the agent on the second-best short-run moral hazard contract***

The expected project output (the resulting effort level), the principal's profit and the paternalism-type joint principal-agent surplus (social welfare) are all eroded by the presence of the agent's short-run impatience and decrease in its degree $(1-\beta)$*, i.e.,* $\frac{\partial e^{SB}_{S[\beta]}}{\partial\beta} > 0$, $\frac{\partial \pi^{SB}_{S[\beta]}}{\partial\beta} = \frac{\partial S^{SB}_{1S[\beta]}}{\partial\beta} > 0$, $\frac{\partial \widetilde{S}^{SB}_{1S[\beta]}}{\partial\beta} = \frac{\partial}{\partial\beta}\left(\pi^{SB}_{S[\beta]} + \tilde{U}^{SB}_{1,P}(\beta)\right) > 0.$

Proof of Proposition 2. In the appendix. ■

The presence of the agent's short-run impatience involves an adverse effect on the principal's profit and social welfare; this constitutes a different finding with DellaVigna-Malmendier (2004), in which the principal's profits and the joint principal-agent surplus are unaffected by the degree of the agent's time-inconsistency $(1-\beta)$ as long as the agent has full self awareness; the underlying reasoning and associated critical difference is: here in our moral hazard model, the agent's short-run impulsive behavior cannot be stimulated or regulated by the instrument in the contract under the second-best information structure (the privacy of the effort level and the associated

second-best contracting environment.

infeasibility of w_1).

In addition, the preceding proposition plus the fact that $\tilde{U}_{1,P}^{SB}(\beta) > U_1^{SB}(\beta) = 0$ indicates: on the one hand, after correcting the individual irrational feeling suffered from the self-control problem by a paternal evaluation, the employee preserves more surplus in the contract; but on the other hand, on the ground of a paternalistic social planner, the joint principal-agent welfare still loses due to the presence of agent's short-run impatience, i.e., in short, the principal sacrifices more than the agent's "gain". This result is also straightforward since the economy results in an inefficient allocation due to the agent's short-run impatience.

The result that the principal bears a loss when encountering a short-run impatient agent naturally raises us a question: according to the (β, δ) logic, in the long-run, the agent is more patient, thus, is the principal interested in advancing the contract time, i.e., the employer starts recruitment from Period 0 but all the tasks take place in Period 1 (3-period long-run employment) instead of the former 2-period one on the sake of taking advantage of long-run patience of the employee (1 according to our normalization) and avoid her short-run impatience (β)? We are going to answer this question in the next section.

3 3-Period (Long-run) Moral Hazard Contract for a Sophisticated Agent

3.1 The Timing Extension

We ask the question: can the principal achieve better outcome and higher profits by advancing the date of signing the contract one period before, i.e., by rearranging a short-run recruitment to a long-run one? Under this kind of consideration, the timeline is extended as in Figure 3. We see that a new Period 0 appears. Compared to the timeline in Section 2, the principal prepares a contract one period in advance with respect to the starting point of work (Period 1). According to the time-inconsistency concept, the utility of the agent's self 0 can be stated as $U_0 = \beta (w_2 - C(e))$.

Figure 3 is about here

This long-run alternative reflects the principal's intention to take advantage of the long-run more patient discounting of the time-inconsistent agent. Our purpose is to see whether this timing rearrangement can alleviate the adverse impact of agent's short-run impatience to the principal's profit and social welfare appeared in the short-run contracting environment.

Because there are altogether 3 periods now, the extent of self-awareness of the agent, i.e., the own prediction about the degree of short-run impatience in next period, begins mattering. For the time being, we assume that the agent is sophisticated, i.e., fully aware of self 1's time preference in Period 0.

And since after signing the contract in Period 0, all the tasks and decisions within the contract take place one period later, the commitment situations of both principal and agent become a question. Is the unilateral deviation from the contract in Period 1 allowed? Does the possibility of renegotiation exist? We should clear out our commitment settings. The principal is assumed to be always committed to fulfill the rights and obligations in the contract which she signed[6]; while for the agent, we separate into 2 cases, where we state in the following 2 definitions.

Definition 5 ***1-sided commitment***

Only the principal can commit to fulfill the contract at all the dates after she signed it, but the agent has the unilateral right to leave the contract freely (without punishment) at each date after signing the contract.

Definition 6 ***2-sided commitment***

Once the contract is signed, all the forthcoming selves of both the principal and the agent are restricted to fulfill all the rights and obligations within the

[6]That is to say: no "ratchet effect". This is consistent with our model setting intention: The principal is assumed to be fully patient, fully rational and fully committed; while all the behavioral bias is on the agent's side, including time preference, self-awareness and lack of commitment.

contract, i.e., no one has possibility to quit the contract unilaterally or propose renegotiation. In other words, the outside option of the impulsive resignation of agent's self 1 is $-\infty$[7].

In fact the difference between these two commitment definitions is only about the restriction toward the agent. We can explain them in this way: in the 2-sided commitment setting, the principal (the contract designer) can constitute an exogenous commitment device to restrict the agent not leaving the contract in the intermediate stages (see also the similar setting in DellaVigna-Malmendier (2004) and Gilpatric (2008)), e.g., considerable punishment for a unilateral escape, etc.; while in the 1-sided commitment, the principal is less powerful to forbid the agent's resignation in Period 1.

3.2 Long-run Contract under 1-sided commitment

Straightforwardly we can infer that if the future unilateral resignation of the agent is not ruled out inside the contract, then a long-run rearrangement is useless since the contract design and the principal still have to be restricted to each period's short-run impatience of the agent. We confirm this point more clearly and detailedly through the following proposition and its proof.

[7]The amount $-\infty$ can be illustrated by an example that it is explicitly stated in the contract: the unilateral repudiation would be put into jail.

Proposition 3 ***Long-run second-best moral hazard contract for a sophisticated time-inconsistent agent under 1-sided commitment***

In this case, the principal cannot do any better than the short-run, i.e., advancing the time of contracting is meaningless.

Proof of Proposition 3. Since the self 0 of the agent is sophisticated, he knows well that his self 1 will choose $e = \beta b$ when facing a payment scheme $w_2 = a + by$ (shown in Section 2.3.2 by solving (10)) if he does not quit unilaterally in Period 1. The principal takes all these into account (backward induction procedure) and derives an optimal contract (a, b). The program is to maximize (11) subject to the participation of the agent's sophisticated self 0 (13), the participation of the agent's self 1 (12) and non-manipulation constraints (6), where the new equation (13) is:

$$U_0 = \beta \left(1 \{U_1 \geq 0\} \cdot \left(a + b \cdot \beta b - \frac{\gamma}{2} b^2 \sigma^2 - \frac{(\beta b)^2}{2} \right) \right) \geq 0 \qquad (13)$$

Since the resulting b is non-negative (left-hand side of non-manipulability constraints), $U_1 \geq 0 \Rightarrow a + b \cdot \beta b - \frac{\gamma}{2} b^2 \sigma^2 \geq 0 \Rightarrow a + b \cdot \beta b - \frac{\gamma}{2} b^2 \sigma^2 > \beta \left(a + b \cdot \beta b - \frac{\gamma}{2} b^2 \sigma^2 \right) \geq \frac{(\beta b)^2}{2}$; therefore $U_1 \geq 0 \Rightarrow U_0 > 0$. So after cancelling the redundant constraint (13), the program coincides exactly with the short-run second-best contract design. ■

Under 1-sided commitment, since the contract offered in the long-run coincides with that in the short-run, the principal cannot do any better by the creation of an advanced date 0. And because there is no difference between the long-run and short-run recruitment with the absence of agent's commitment, not only the employer's profit, but also all the other economic outcomes (including expected project output, paternalism-type consumer/agent surplus and joint principal-agent welfare) stay at the same levels as in the short-run.

3.3 Long-run Contract under 2-sided Commitment

On the behalf of the principal, the difference introduced by 2-sided commitment instead of 1-sided one is the elimination of constraint (12) (the participation of the agent's self 1) and the simplification of constraint (13) to (14) about the participation of the agent's self 0:

$$U_0 = \beta \left(a + b \cdot \beta b - \frac{\gamma}{2} b^2 \sigma^2 - \frac{(\beta b)^2}{2} \right) \geq 0 \tag{14}$$

The principal maximizes (11) subject to (6)(14), which gives us the long-run second-best contract under 2-sided commitment and its associated economic outcomes as:

$$\begin{cases} b^{SB}_{L[\beta]2} = \min\left\{\frac{\beta}{\beta^2+\gamma\sigma^2}, 1\right\} \\ a^{SB}_{L[\beta]2} = \frac{1}{2}\left(b^{SB}_{L[\beta]2}\right)^2(\gamma\sigma^2 + \beta^2 - 2\beta) \\ e^{SB}_{L[\beta]2} = \min\left\{\frac{\beta^2}{\beta^2+\gamma\sigma^2}, \beta\right\} \\ \pi^{SB}_{L[\beta]2} = S^{SB}_{0L[\beta]2} = \begin{cases} \frac{\beta^2}{2(\beta^2+\gamma\sigma^2)}, \text{ when } b^{SB}_{L[\beta]2} = \frac{\beta}{\beta^2+\gamma\sigma^2} \\ -a^{SB}_{L[\beta]2} = \beta - \frac{\beta^2}{2} - \frac{\gamma}{2}\sigma^2, \text{ when } b^{SB}_{L[\beta]2} = 1 \end{cases} \end{cases}$$

where the subscript "L" stands for "long-run" and the "2" after "$[\beta]$" stands for "2-sided commitment". The detailed solving procedure is in the appendix. From the following Proposition 4 and its proof, both the principal's profit and all the other economic outcomes will get strict amelioration than their short-run counterparts under this 2-sided commitment environment, so "contracting earlier" happens in this case.

Proposition 4 ***Long-run second-best moral hazard contract for a sophisticated time-inconsistent agent under 2-sided commitment***

(i) The principal's profit, the expected project output (the agent's effort level) and the joint principal-agent welfare are all higher than the counterparts in the short-run environment; but on the other hand, all the indices are lower than the ones in the contract with a time-consistent/fully patient agent: Mathematically, $e^{SB}_{S[\beta]} \leq e^{SB}_{L[\beta]2} < e^{SB}_{[1]}$, $\pi^{SB}_{S[\beta]} < \pi^{SB}_{L[\beta]2} < \pi^{SB}_{[1]}$, $S^{SB}_{0S[\beta]} = S^{SB}_{1S[\beta]} \leq S^{SB}_{0L[\beta]2} < S^{SB}_{0[1]} = S^{SB}_{1[1]}$, *where the two equalities hold if the agent is risk-neutral* ($\gamma = 0$).

(ii) The moral hazard problem still exists, i.e., the second-best expected project output, the profit and the joint principal-agent welfare are all still inferior to their corresponding first-best levels, even if the agent is risk-neutral: $e_{L[\beta]2}^{SB}(\gamma = 0) < e^{FB} = 1$, $\pi_{L[\beta]2}^{SB}(\gamma = 0) < \pi^{FB} = \frac{1}{2}$, $S_{0L[\beta]2}^{SB}(\gamma = 0) < S_0^{FB} = S_1^{FB} = \frac{1}{2}$.

(iii) When the agent is not very risk-averse ($\gamma < \frac{1}{4\sigma^2}$), the upper-bound non-manipulability constraint $b \leq 1$ may be binding, more specifically, $b_{L[\beta]2}^{SB}$ binds at 1 *when the parameters $(\beta, \gamma, \sigma^2)$ locate at the interior area of the set* $\Gamma = \left\{(\beta, \gamma, \sigma^2) \mid \sigma^2 > 0, \gamma \in \left[0, \frac{1}{4\sigma^2}\right], \beta \in \left[\frac{1}{2} - \sqrt{\frac{1}{4} - \gamma\sigma^2}, \frac{1}{2} + \sqrt{\frac{1}{4} - \gamma\sigma^2}\right]\right\}$.

Proof of Proposition 4. In the appendix. ■

The amelioration of the principal and the society in the long-run contracting environment under the 2-sided commitment thanks to the prevention of the short-run self's impulsive resignation. By observing the participation constraint (14), the β outside the bracket directly cancels, so the maximization program is just like contracting with a time-consistent/fully patient agent; from here the original purpose of taking advantage of the agent's long-run patience is achieved. On the other hand, the fact that the adverse impact of the agent's time-inconsistency/short-run impatience cannot be fully alleviated is also straightforward. Both the principal and the long-run self of the agent have no means to restrict/stimulate the short-run agent self's own choice of private effort level: this constitutes the inferiority of optimal long-run second-best time-inconsistent agent's contract than that of a time-consistent agent's case.

Points (ii)(iii) in Proposition 4 exhibit the violations to the two standard moral hazard properties established in the late 1970s (Harris-Raviv, 1976; Holmstrom, 1979) which are stated in Proposition 1: this expresses the important impact of the agent's time-inconsistency to the moral hazard contract framework.

Furthermore, a "countervailing" financial phenomenon may happen when the agent is time-inconsistent and not very risk-averse; specifically, $b = 1$ represents the fact that the principal sells the whole project to the agent and just gets a fixed dividend $(-a^{SB}_{L[\beta]2})$ (it is not difficult to check that when $b^{SB}_{L[\beta]2} = 1$, $a^{SB}_{L[\beta]2}$ is always negative). And with referring to the appendix, we can rewrite the solution of $b^{SB}_{L[\beta]2}$ as:

$$b^{SB}_{L[\beta]2} \begin{cases} < 1, \text{ when } (\beta, \gamma, \sigma^2) \notin \Gamma \\ = 1, \text{ when } (\beta, \gamma, \sigma^2) \in \Gamma \end{cases}$$

in fact Γ represents the set of parameters such that the full delegation (the selling behavior) happens.

And in addition, we notice that if $\gamma = 0$ (the agent is also risk-neutral as the principal), the interval for $\beta \in \Gamma$ expands to $[0, 1]$, which means that the risk-neutral principal will always totally delegate to the time-inconsistent risk-neutral agent, i.e., the risk-neutral principal here behaves like a fully insured risk-averse investor under the presence of the time-inconsistency plus risk-neutrality of the agent.

Figure 4 illustrates the area of Γ (the shadowed area) where the principal sells the whole project to the sophisticated agent; we notice that this happens only when the agent is not too risk-averse ($\gamma \leq \frac{1}{4\sigma^2}$); if the risk-averse extent exceeds this threshold, the full delegation cannot be sustained for $\forall \beta$; and when $\gamma \leq \frac{1}{4\sigma^2}$, the vertical measure of Γ set (the interval of β to sell the project to the agent given a risk tolerance $\frac{1}{\gamma}$) shrinks with higher γ (more risk averse). When $b < 1$, the principal is still like an employer who hires a worker and shares the risk of the project with the agent; b is the incentive component inside the contract which assures that the highest possible effort level in Period 1 can be exerted by the agent's self 1.

Figure 4 is about here

These findings convey an interesting and important point: the time-inconsistency can be regarded as a complement to the risk tolerance in this context.

4 The Impact of Agent's Naivety on the Long-run Moral Hazard Contract under 2-sided commitment

In the long-run (3-period) timeline, we clear out the concept of naivety/sophistication according to the introduction and the literature.

Definition 7 ***(O'Donoghue-Rabin (2001)) Naivety/Sophistication***

$\widehat{\beta}$ ($\beta \leq \widehat{\beta} \leq 1$) is denoted as the subjective prediction by the agent's self 0 toward his self 1's short-run impatience level.

(i) $\widehat{\beta} = \beta$: the agent is called sophisticated.

(ii) $\widehat{\beta} = 1$: the agent is called fully naive.

(iii) $\beta < \widehat{\beta} < 1$: the agent is called partially naive.

We see that only over-confident bias ($\widehat{\beta} > \beta$) is considered. Based on my knowledge, no literature deals with the over-pessimism prediction toward self-control ($\widehat{\beta} < \beta$)[8].

Firstly, the agent's naivety only matters in the 3-period contracting environment. Secondly, if with 1-sided commitment, the presence of the agent's naivety will alter nothing compared to previous sections since only the agent's self 1 is the contracting object of the principal; although the self 0's naive

[8] In the literature, $\widehat{\beta} < \beta$ is not consistent with the experimental and field evidence.

participation and over-confident effort exertion is asserted, the rational principal will not be confused by this; the real contract scheme and the associated outcomes, conclusions will not be influenced. So we restrict our concentration toward naivety on the 3-period 2-sided commitment contract.

The objective of the principal (11) and the non-manipulability constraints (6) do not change; only the self 0's participation based on his prediction toward tomorrow's behavior has a difference. We still solve the optimal contract by backward induction. Facing a contract form (a, b), a naive self 0 predicts that his self 1 would choose effort based on the following maximization:

$$\max_e \widehat{U}_1 = \widehat{\beta}\left(a + be - \frac{\gamma}{2}b^2\sigma^2\right) - \frac{e^2}{2} \tag{15}$$

where the solution is $\widetilde{e} = \widehat{\beta}b$. And then the self 0 inserted $\widetilde{e}$ into her evaluation and constitutes the participation constraint:

$$\widehat{U}_0 = \beta\left(a + b \cdot \widehat{\beta}b - \frac{\gamma}{2}b^2\sigma^2 - \frac{\left(\widehat{\beta}b\right)^2}{2}\right) \geq 0 \tag{16}$$

The principal takes all these into account, i.e., (15)(16) and maximizes (11) further subject to (6). We get the optimal contract as follows (the detailed solving procedure is in the appendix):

$$\begin{cases} \widehat{b}_{L[\beta]2}^{SB} = \begin{cases} \frac{\beta}{\widehat{\beta}^2+\gamma\sigma^2-2\left(\widehat{\beta}-\beta\right)}, \text{ if } \gamma\sigma^2 \geq \widehat{\beta}\left(1-\widehat{\beta}\right)+\left(\widehat{\beta}-\beta\right) \\ 1, \text{ otherwise} \end{cases} \\ \widehat{a}_{L[\beta]2}^{SB} = \frac{1}{2}\left(\widehat{b}_{L[\beta]2}^{SB}\right)^2\left(\gamma\sigma^2+\widehat{\beta}^2-2\widehat{\beta}\right) \\ \widehat{e}_{L[\beta]2}^{SB} = \begin{cases} \frac{\beta^2}{\widehat{\beta}^2+\gamma\sigma^2-2\left(\widehat{\beta}-\beta\right)}, \text{ if } \widehat{b}_{L[\beta]2}^{SB} = \frac{\beta}{\widehat{\beta}^2+\gamma\sigma^2-2\left(\widehat{\beta}-\beta\right)} \\ \beta, \text{ if } \widehat{b}_{L[\beta]2}^{SB} = 1 \end{cases} \end{cases}$$

where the additional symbol $\widehat{}$ represents outcomes for a naive agent. We notice that here $\widehat{e}_{L[\beta]2}^{SB}$ denotes the real effort level exerted by self 1 ($= \beta\widehat{b}_{L[\beta]2}^{SB}$), while not the naively predicted effort level by self 0 ($\widehat{\beta\widehat{b}_{L[\beta]2}^{SB}}$). The following proposition states the impact of the agent's naivety on the principal's project outcome and the profit.

Proposition 5 ***The impact of the agent's naivety on the principal***

The following two relationships hold:

(i) $e_{L[\beta]2}^{SB} \leq \widehat{e}_{L[\beta]2}^{SB} < e_{[1]}^{SB}$*: the expected project output under the agent's naivety is higher than the sophisticated agent's case, but any time-inconsistent agent (no matter how naive or sophisticated) works less diligently than a time-consistent agent.*

(ii) $\pi_{L[\beta]2}^{SB} < \widehat{\pi}_{L[\beta]2}^{SB} < \pi_{[1]}^{SB}$*: the principal's profit is strictly better off by the presence of agent's naivety but is still inferior to the level of a time-consistent agent.*

Proof of Proposition 5. In the appendix. ■

Thanks to the agent's imperfect self-awareness, the principal can realize higher project output and extract more profit. The intuition is straightforward since the extent of the naivety relaxes the participation constraint of self 0 and benefits the rational principal, which is consistent with the finding in DellaVigna-Malmendier (2004) on the consumer's naivety. But the principal cannot return back the time-consistent agent's world since there exists no instrument in the contract which can stimulate the self 1's short-run harder working.

On the side of the agent, the true realized utility of self 0 is

$$\begin{aligned}
\widehat{U}^{SB}_{0L[\beta]2} &= \beta\left(\widehat{a}^{SB}_{L[\beta]2} + \widehat{b}^{SB}_{L[\beta]2}\widehat{e}^{SB}_{L[\beta]2} - \frac{\gamma}{2}\left(\widehat{b}^{SB}_{L[\beta]2}\right)^2\sigma^2 - \frac{\left(\widehat{e}^{SB}_{L[\beta]2}\right)^2}{2}\right) \\
&= \beta\left(\left(\beta - \frac{\beta^2}{2}\right) - \left(\widehat{\beta} - \frac{\widehat{\beta}^2}{2}\right)\right)\left(\widehat{b}^{SB}_{L[\beta]2}\right)^2,
\end{aligned}$$

while not $\widehat{U}_0 = 0$ (the fictitious utility based on naive prediction). It's not difficult to see that $\widehat{U}^{SB}_{0L[\beta]2} < 0$ as long as $\widehat{\beta} > \beta$ since the naive self 0 is cheated by the rational principal to some extent.

On top of the principal and the agent, what should be the social preference toward naivety on the behalf of two types' social planner respectively? And will they agree with each other? The following proposition gives an interesting exhibition.

Proposition 6 ***The impact of the agent's naivety on the society***

(i) If the social planner is paternalism type, then $\widetilde{\widehat{S}}^{SB}_{0L[\beta]2,P} \leq \widetilde{S}^{SB}_{0L[\beta]2,P}$, *where* $\widetilde{\widehat{S}}^{SB}_{0L[\beta]2,P} = \widehat{\pi}^{SB}_{L[\beta]2} + \widetilde{\widehat{U}}^{SB}_{0L[\beta]2} = \widehat{\pi}^{SB}_{L[\beta]2} + \frac{1}{\beta}\widehat{U}^{SB}_{0L[\beta]2}$.

(ii) But if the social planner is let-it-be type, then $\widehat{S}^{SB}_{0L[\beta]2} > S^{SB}_{0L[\beta]2}$ *if* $(\beta, \gamma, \sigma^2) \in \Gamma$, *where* $\widehat{S}^{SB}_{0L[\beta]2} = \widehat{\pi}^{SB}_{L[\beta]2} + \widehat{U}^{SB}_{0L[\beta]2}$.

Proof of Proposition 6. In the appendix. ∎

On the point of view of a paternalistic social planner, although on the one hand the principal is benefited from the agent's naivety, the joint principal-agent paternalism-type surplus still deteriorates, thus the naivety results in a paternalism-type over-production ($\widehat{e}^{SB}_{L[\beta]2} \geq e^{SB}_{L[\beta]2}$ always holds). This conveys similar result on the evaluation toward naivety as DellaVigna-Malmendier (2004) when they adopt the paternalism-type consumer surplus proposed by O'Donoghue-Rabin (2001).

But if the evaluation rationale of the social planner is let-it-be, the social preference toward naivety is ambiguous, which depends on the extent of naivety and the comparison between the agent's short-run impatience and his risk tolerance. A sufficient but not necessary condition for a favorable naivety is: If originally the principal is willing to sell the project to the sophisticated agent ($(\beta, \gamma, \sigma^2)$ exhibits a "super risk-neutrality" combination, see Figure 4 and (iii) of Proposition 4), then, when the agent becomes naive, the joint let-it-be welfare is always benefited from the naivety, though one party (the agent) bears a loss in its own utility ($\widehat{U}^{SB}_{0L[\beta]2} < 0$). So the naivety leaves a caveat of the way to define consumer surplus in our context.

At last, is the complementary relationship between short-run impatience and risk attitude found in Section 3.3 altered by the presence of the agent's naivety? Does it vanish, being attenuated or being exaggerated? Before answering, we define

$$\Omega \triangleq \left\{ \left(\beta, \widehat{\beta}, \gamma, \sigma^2\right) \middle| \sigma^2 > 0, \gamma \geq 0, 0 < \beta \leq \widehat{\beta} \leq 1, \widehat{b}^{SB}_{L[\beta]2} = 1 \right\},$$

i.e., the parameters set which leads to full delegation when a risk-neutral principal faces a risk-averse, naive and time-inconsistent agent. We can see the preceding Γ set is just a particular example of this general Ω where $\Gamma = \Omega\left(\widehat{\beta} = \beta\right)$. Based on the definition of Ω, the following proposition clears out a definitive answer that the presence of naivety strengthens the complementarity between $(1 - \beta)$ and γ.

Proposition 7 ***The willingness of the principal to sell the project to a sophisticated agent or a naive agent***

$\forall \beta, \widehat{\beta}, \gamma, \sigma^2$ *such that* $\sigma^2 > 0, \gamma \geq 0, 0 < \beta \leq \widehat{\beta} \leq 1$, *we have* $\Gamma \subseteq \Omega$.

Proof of Proposition 7. In the appendix. ■

Proposition 7 tells an interesting outcome: if the principal is willing to sell the project to a sophisticated agent, then the "selling" behavior will also definitively happen when the agent becomes naive; but conversely it is not necessarily true. And Figure 5 gives an illustration and comparison between sophisticated contract and fully naive contract. Γ and $\Omega\left(\widehat{\beta}=1\right)$ represent the two extremes of the evolution of Ω from $\widehat{\beta}=\beta$ to $\widehat{\beta}=1$. We see that the full naivety expands the probability of full delegation occurrence and exhibits a perfect linear complementarity between time preference and risk attitude.

Figure 5 is about here

5 Concluding remarks

Our model starts from a story in which a time-consistent risk-neutral principal is hiring a short-run impatient and risk-averse agent to help fulfill a risky project where the production procedure involves an exogenous uncertainty and is the agent's private information, so moral hazard possibly arises. To stimulate the agent's autonomous working, the principal employs a linear incentive scheme which depends on the final output handed in by the agent.

First we apparently see the negative impact of the agent's short-run impatience on the project, the principal's profits and the social welfare. Then we ask a question whether the principal can do better by rearranging such a short-run recruitment relationship to a long-run one since the time-inconsistency resulted from short-run impatience exhibits a long-run patience. In a long-run contract, each party's unilateral commitment situation matters substantially. We restrict us to the case of no "ratchet effect", i.e., the principal is always assumed to commit herself perfectly. As for the agent, we divide into 2 cases: 1-sided commitment, under which the agent cannot commit his future self's participation; 2-sided commitment, under which both parties inside the contract commit their future selves to fulfill the obligations signed beforehand, i.e., there exists some external (principal's punishment power) or internal commitment device to prevent the agent future self's impulsive resignation. We find that under the 2-sided commitment environment, the principal can obtain higher profits in a long-run employment relationship than a short-run hiring. And the agent may also regard the exogenous commitment device (by the principal or the society) as a self-discipline instrument to partially correct the self control problem.

In some recent literatures concerning time-inconsistent agent's adverse selection problem (e.g., Gilpatric (2008)), the negative impact of the agent's self control defect can be entirely prevented in a full commitment environment for all parties; but this is not true in our moral hazard setting, i.e., all economic outcomes including the expected project output, the profit, the social welfare, etc., are inferior to that of a time-consistent agent even under full commitment restrictions for both players. A straightforward reason is that under the setting of moral hazard, the short-run private effort level cannot be stimulated by neither the principal nor the agent's long-run patient self.

According to our findings under the full commitment environment of a sophisticated time-inconsistent agent's long-run moral hazard problem, two violations to traditional moral hazard properties established in the late 1970s (1. equivalence between first-best and second-best under agent's risk-neutrality; 2. slackness of non-manipulability constraints) shed some lights on the essential nature of the concept of time-inconsistency: here the extent of short-run impatience can be regarded as a complement of risk tolerance. If the agent holds a reasonable (not too much) risk-aversion and combined with some degree of time-inconsistency, then the risk-neutral principal believes that this agent is more appropriate to manage the risky asset than herself and is willing to sell the whole project to the agent, while the risk-neutral principal herself just gets a fixed dividend as a seemingly risk-averse investor.

At last, the lack of complete self-awareness of the agent toward his future impulsion is always good news to the principal, but its impact on the social welfare is perplexing, which depends on the way of defining consumer surplus for a naive agent. The complementary relationship between time-inconsistency/short-run impatience and risk tolerance is strengthened and exaggerated by the presence of the agent's naivety. A fully naive agent exhibits a perfect linear complementarity between time preference and risk attitude.

After these, some aspects of extensions may be expected.

First, besides the privacy of the agent's effort level (source of moral hazard), the principal may also have incomplete information towards some characterizations of the agent such as the degree of short-run impatience β, the production ability, etc. This constitutes a joint adverse selection and moral hazard problem with a time-inconsistent agent in which no literature works out general insights yet to my knowledge.

Second, according to latest finding like Kylymnyuk (2009), the cooperation among time-inconsistent agents may promote efficiency, so two agents' joint work within the production procedure (e.g., moral hazard in team work) is also interesting to investigate.

Then, on the aspect of empirics, we should resort to some field and laboratory studies to verify the existence of individual's short-run impatience/time-inconsistency in contracting by observing the abnormal financial arrangement of reselling a risky asset from a risk-neutral principal to a risk-averse agent in the reality or experiments. Furthermore, by observing different moral hazard contracting schemes among different countries/jobs, we may infer the different extents of short-run impatience/time-inconsistency among different economic environments and nationalities.

Appendix

Proof of Proposition 2. $e^{SB}_{S[\beta]} = \frac{\beta^2}{\beta+\gamma\sigma^2} = \frac{1}{\frac{1}{\beta}+\frac{\gamma\sigma^2}{\beta^2}}$, where the denominator is decreasing in β, so $\frac{\partial e^{SB}_{S[\beta]}}{\partial\beta} > 0$.

The same result holds for the profit since $\pi^{SB}_{S[\beta]} = \frac{1}{2}e^{SB}_{S[\beta]}$.

$\tilde{U}^{SB}_{1,P}(\beta) = \frac{\beta^3(1-\beta)}{2(\beta+\gamma\sigma^2)^2}$, so $S^{SB}_{1S[\beta]} = \pi^{SB}_{S[\beta]} + \tilde{U}^{SB}_{1,P}(\beta) = \frac{2\beta^3+\beta^2\left(\gamma\sigma^2-\beta^2\right)}{2(\beta+\gamma\sigma^2)^2} = \frac{\frac{2}{\beta}+\frac{\gamma\sigma^2}{\beta^2}-1}{2\left(\frac{1}{\beta}+\frac{\gamma\sigma^2}{\beta^2}\right)^2} = \frac{2\zeta+\Lambda\zeta^2-1}{2(\zeta+\Lambda\zeta^2)^2}$, where $\zeta = \frac{1}{\beta}$ and $\Lambda = \gamma\sigma^2$. Thus $\frac{\partial S^{SB}_{1S[\beta]}}{\partial\zeta} = \frac{(2+2\Lambda\zeta)2\left(\zeta+\Lambda\zeta^2\right)^2-\left(2\zeta+\Lambda\zeta^2-1\right)4\left(\zeta+\Lambda\zeta^2\right)^2(1+2\Lambda\zeta)}{4(\zeta+\Lambda\zeta^2)^4} = \frac{(1+\Lambda\zeta)\left(\zeta+\Lambda\zeta^2\right)-\left(2\zeta+\Lambda\zeta^2-1\right)(1+2\Lambda\zeta)}{(\zeta+\Lambda\zeta^2)^3} = \frac{\zeta(1+\Lambda\zeta)^2-2\zeta-4\Lambda\zeta^2-\Lambda\zeta^2-2\Lambda^2\zeta^3+1+2\Lambda\zeta}{(\zeta+\Lambda\zeta^2)^3} = \frac{1-\zeta+2\Lambda\zeta-3\Lambda\zeta^2-\Lambda^2\zeta^3}{(\zeta+\Lambda\zeta^2)^3} = \frac{(1-\zeta)(1+2\Lambda\zeta)-\Lambda\zeta^2-\Lambda^2\zeta^3}{(\zeta+\Lambda\zeta^2)^3} < 0$, where the last inequality comes from $1-\zeta = 1-\frac{1}{\beta} < 0$. So $S^{SB}_{1S[\beta]}$ is decreasing in ζ, therefore is increasing in β, i.e., $\frac{\partial S^{SB}_{1S[\beta]}}{\partial\beta} > 0$. ■

The solution of maximizing (11) subject to (14)(6) and the associated economic outcomes.

At first we ignore the non-manipulability constraints.

(14) has to be binding so that $a = \frac{1}{2}b^2\left(\gamma\sigma^2+\beta^2-2\beta\right)$, then (11) becomes an unconstrained maximization over b. The solution is $\tilde{b} = \frac{\beta}{\gamma\sigma^2+\beta^2} > 0$. But we can see that $\tilde{b}$ may exceed 1 for some β, γ and σ. We further write out the condition of $\tilde{b} > 1$, i.e., where the non-manipulability constraint is going to be binding: $\gamma < \frac{1}{4\sigma^2}$ and $\frac{1}{2}-\sqrt{\frac{1}{4}-\gamma\sigma^2} < \beta < \frac{1}{2}+\sqrt{\frac{1}{4}-\gamma\sigma^2}$; or we can express the set of parameters β, γ and σ such that $b = 1$ as: $\Gamma = \left\{(\beta,\gamma,\sigma^2) \mid \sigma^2 > 0, \gamma \in \left[0, \frac{1}{4\sigma^2}\right], \beta \in \left[\frac{1}{2}-\sqrt{\frac{1}{4}-\gamma\sigma^2}, \frac{1}{2}+\sqrt{\frac{1}{4}-\gamma\sigma^2}\right]\right\}$.

So the complete solution is: $b^* \begin{cases} = 1, \text{ when } (\beta,\gamma,\sigma^2) \in \Gamma \\ = \frac{\beta}{\gamma\sigma^2+\beta^2}, \text{ when } (\beta,\gamma,\sigma^2) \notin \Gamma \end{cases}$. Or we express it more compactly: $b^* = \min\left\{\frac{\beta}{\gamma\sigma^2+\beta^2}, 1\right\}$. Consequently $a^* = \frac{1}{2}(b^*)^2(\gamma\sigma^2+\beta^2-2\beta)$.

$e^{SB}_{L[\beta]2} = \beta b^*$ (the optimal behavior of the agent's self 1 by solving (10)).

$\pi^{SB}_{L[\beta]2} = S^{SB}_{0L[\beta]2}$ since the participation constraint of the agent's self 0 is binding, and its value can be obtained by imposing separately for $b^* = 1$ and $b^* = \frac{\beta}{\gamma\sigma^2+\beta^2}$ into the maximized program (11) of the principal. ■

Proof of Proposition 4.

(i)

About $e^{SB}_{S[\beta]} \leq e^{SB}_{L[\beta]2} < e^{SB}_{[1]}$, where the equality holds only if $\gamma = 0$:

$e^{SB}_{S[\beta]} = \frac{\beta^2}{\beta+\gamma\sigma^2}$, $e^{SB}_{L[\beta]2} = \min\left\{\frac{\beta^2}{\beta^2+\gamma\sigma^2}, \beta\right\}$, $e^{SB}_{[1]} = \frac{1}{1+\gamma\sigma^2}$

$e^{SB}_{L[\beta]2} < e^{SB}_{[1]}$: $\min\left\{\frac{\beta^2}{\beta^2+\gamma\sigma^2}, \beta\right\} \leq \frac{\beta^2}{\beta^2+\gamma\sigma^2} = \frac{1}{1+\frac{\gamma\sigma^2}{\beta^2}} < \frac{1}{1+\gamma\sigma^2}$ since $\beta < 1$.

$e^{SB}_{S[\beta]} < e^{SB}_{L[\beta]2}$ when $\gamma > 0$: $\beta < 1 \Rightarrow \frac{\beta^2}{\beta+\gamma\sigma^2} < \frac{\beta^2}{\beta^2+\gamma\sigma^2}$; $\gamma > 0 \Rightarrow \beta^2 < \beta^2 + \beta\gamma\sigma^2 \Rightarrow \frac{\beta^2}{\beta+\gamma\sigma^2} < \beta$; so altogether $\frac{\beta^2}{\beta+\gamma\sigma^2} < \min\left\{\frac{\beta^2}{\beta^2+\gamma\sigma^2}, \beta\right\}$.

$e^{SB}_{S[\beta]} = e^{SB}_{L[\beta]2}$ when $\gamma = 0$: by the expression of $e^{SB}_{S[\beta]}$ and $e^{SB}_{L[\beta]2}$, $\gamma = 0 \Rightarrow e^{SB}_{L[\beta]2} = e^{SB}_{S[\beta]} = \beta$.

About $\pi^{SB}_{S[\beta]} < \pi^{SB}_{L[\beta]2} < \pi^{SB}_{[1]}$, $S^{SB}_{1S[\beta]} < \widetilde{S}^{SB}_{1S[\beta],P} \leq S^{SB}_{0L[\beta]2} = \widetilde{S}^{SB}_{0L[\beta]2,P} < S^{SB}_{0[1]} = S^{SB}_{1[1]}$, where the equality holds only if $\gamma = 0$:

$\pi^{SB}_{S[\beta]} = S^{SB}_{1S[\beta]} < \widetilde{S}^{SB}_{1S[\beta],P}$ is obvious since $\tilde{U}^{SB}_{1,P}(\beta) > 0 = U^{SB}_1(\beta)$. Plus the facts that $S^{SB}_{0L[\beta]2} = \widetilde{S}^{SB}_{0L[\beta]2,P} = \pi^{SB}_{L[\beta]2}$ (under long-run 2-sided commitment, the two types of social welfare coincide since $\tilde{U}^{SB}_{0L2,P}(\beta) = \frac{1}{\beta}U^{SB}_{0L2}(\beta) = 0$) and $S^{SB}_{[1]} = \pi^{SB}_{[1]}$, it suffices to prove the preceding relationships among π and S by the following simplified inequalities: $\pi^{SB}_{[1]} > \pi^{SB}_{L[\beta]2} \geq \widetilde{S}^{SB}_{1S[\beta],P}$, where the equality holds only if $\gamma = 0$.

$\pi^{SB}_{[1]} > \pi^{SB}_{L[\beta]2}$:

By the preceding solving procedure of maximizing (11) subject to (14)(6), we know that if ignoring (6), the unconstrained solution for b is $\tilde{b} = \frac{\beta}{\gamma\sigma^2+\beta^2}$, which generates an unconstrained profit $\tilde{\pi} = \frac{\beta^2}{2(\beta^2+\gamma\sigma^2)}$. But when (6) matters, i.e., b^* binds at 1, the profit becomes $\beta - \frac{\beta^2}{2} - \frac{\gamma}{2}\sigma^2$. By the principle of maximization, $\beta - \frac{\beta^2}{2} - \frac{\gamma}{2}\sigma^2 \leq \frac{\beta^2}{2(\beta^2+\gamma\sigma^2)}$ when $b^{SB}_{L[\beta]2} = 1$. So no matter $b^{SB}_{L[\beta]2} < 1$ or $= 1$, $\pi^{SB}_{L[\beta]2} \leq \frac{\beta^2}{2(\beta^2+\gamma\sigma^2)} < \frac{1}{2(1+\gamma\sigma^2)} = \pi^{SB}_{[1]}$.

$\pi^{SB}_{L[\beta]2} \geq \widetilde{S}^{SB}_{1S[\beta],P}$, where the equality holds if $\gamma = 0$:

$\gamma = 0 \Rightarrow b^{SB}_{L[\beta]2} = 1$, $\pi^{SB}_{L[\beta]2} = \beta - \frac{\beta^2}{2}$, $\pi^{SB}_{S[\beta]} = \frac{\beta}{2}$, $\tilde{U}^{SB}_{1,P}(\beta) = \frac{\beta(1-\beta)}{2}$, obviously $\pi^{SB}_{L[\beta]2} = \frac{\beta}{2} + \frac{\beta(1-\beta)}{2} = \pi^{SB}_{S[\beta]} + \tilde{U}^{SB}_{1,P}(\beta) = \widetilde{S}^{SB}_{1S[\beta],P}$.

$\gamma > 0$: When the constraint "$b \leq 1$" does not bind, i.e., $b^{SB}_{L[\beta]2} = \frac{\beta}{\gamma\sigma^2+\beta^2} \leq 1$, $\pi^{SB}_{L[\beta]2} - \widetilde{S}^{SB}_{1S[\beta],P} = \pi^{SB}_{L[\beta]2} - \pi^{SB}_{S[\beta]} - \tilde{U}^{SB}_{1,P}(\beta) = \frac{1}{2}\left(\frac{\beta^2}{\beta^2+\gamma\sigma^2} - \frac{\beta^2}{\beta+\gamma\sigma^2} - \frac{\beta^3(1-\beta)}{(\beta+\gamma\sigma^2)^2}\right) = \frac{1}{2}\left(\frac{\beta^3(1-\beta)}{(\beta^2+\gamma\sigma^2)(\beta+\gamma\sigma^2)} - \frac{\beta^3(1-\beta)}{(\beta+\gamma\sigma^2)^2}\right) > \frac{1}{2}\left(\frac{\beta^3(1-\beta)}{(\beta+\gamma\sigma^2)(\beta+\gamma\sigma^2)} - \frac{\beta^3(1-\beta)}{(\beta+\gamma\sigma^2)^2}\right) = 0$; when $b^{SB}_{L[\beta]2}$ binds at 1, i.e., $\frac{\beta}{\gamma\sigma^2+\beta^2} > 1$, $\pi^{SB}_{L[\beta]2} - \widetilde{S}^{SB}_{1S[\beta],P} = \pi^{SB}_{L[\beta]2} - \pi^{SB}_{S[\beta]} - \tilde{U}^{SB}_{1,P}(\beta) = \frac{1}{2}\left((2\beta - \beta^2 - \gamma\sigma^2) - \frac{\beta^2}{\beta+\gamma\sigma^2} - \frac{\beta^3(1-\beta)}{(\beta+\gamma\sigma^2)^2}\right) = \frac{1}{2}\left(\frac{\beta^2-\beta^3+\gamma\sigma^2(\beta-\beta^2-\gamma\sigma^2)}{\beta+\gamma\sigma^2} - \frac{\beta^3(1-\beta)}{(\beta+\gamma\sigma^2)^2}\right) >$

$\frac{1}{2}\left(\frac{\beta^2(1-\beta)}{\beta+\gamma\sigma^2}-\frac{\beta^3(1-\beta)}{(\beta+\gamma\sigma^2)^2}\right)=\frac{\beta^2(1-\beta)}{2(\beta+\gamma\sigma^2)}\left(1-\frac{\beta}{\beta+\gamma\sigma^2}\right)>0$, where the first inequality comes from the fact that $\frac{\beta}{\gamma\sigma^2+\beta^2}>1$.

(ii)

$e^{SB}_{L[\beta]2}(\gamma=0)=\beta<1$.

$\pi^{SB}_{L[\beta]2}(\gamma=0)=S^{SB}_{0L[\beta]2}(\gamma=0)=\beta-\frac{\beta^2}{2}\triangleq f(\beta)$. $f'(\beta)=1-\beta$, $f'(\beta)=-1<0$. So $f(\beta)$ attains its unique global maximum at $\beta=1$ for $f(1)=\frac{1}{2}$. Therefore, $\beta<1\Rightarrow f(\beta)<\frac{1}{2}$.

(iii)

Refer to the preceding title "The solution of maximizing (11) subject to (14)(6) and the associated economic outcomes". ■

The solution of maximizing (11) subject to (16)(6).

After substituting a from imposing the equality of (16), we get:

$$\max_b \beta b-\frac{1}{2}\left(\gamma\sigma^2+\widehat{\beta}^2-2\left(\widehat{\beta}-\beta\right)\right)b^2. \tag{17}$$

The FOC free from the non-manipulability constraints gives $b^*=\frac{\beta}{\gamma\sigma^2+\widehat{\beta}^2-2\left(\widehat{\beta}-\beta\right)}$. Then we discuss in different cases.

Case 1. $\gamma\sigma^2+\widehat{\beta}^2-2\left(\widehat{\beta}-\beta\right)<0\Leftrightarrow\gamma\sigma^2+\beta^2<\widehat{\beta}\left(1-\widehat{\beta}\right)+\left(\widehat{\beta}-\beta\right)$

The objective function (17) is convex and so $b^*<0$ is minimum, thus the solution in the interval $[0,1]$ should be $b=1$.

Case 2. $\gamma\sigma^2+\widehat{\beta}^2-2\left(\widehat{\beta}-\beta\right)=0\Leftrightarrow\gamma\sigma^2+\beta^2=\widehat{\beta}\left(1-\widehat{\beta}\right)+\left(\widehat{\beta}-\beta\right)$

(17) is linear. Since $\beta>0$, the solution in the interval $[0,1]$ should be

$b = 1$.

Case 3. $0 < \gamma\sigma^2 + \widehat{\beta}^2 - 2\left(\widehat{\beta} - \beta\right) < \beta \Leftrightarrow \gamma\sigma^2 < \widehat{\beta}\left(1 - \widehat{\beta}\right) + \left(\widehat{\beta} - \beta\right) < \gamma\sigma^2 + \beta^2$

(17) is concave and $b^* > 1$ is the global maximum. So the solution in the interval $[0, 1]$ should be $b = 1$.

Case 4. $\gamma\sigma^2 + \widehat{\beta}^2 - 2\left(\widehat{\beta} - \beta\right) \geq \beta \Leftrightarrow \gamma\sigma^2 \geq \widehat{\beta}\left(1 - \widehat{\beta}\right) + \left(\widehat{\beta} - \beta\right)$

(17) is concave and $b^* \in (0, 1]$ is the global maximum. So the solution is just b^*.

We combine Case 1-3 and so the two expressions for $\widehat{b}^{SB}_{L[\beta]2}$ come out. Expressions for $\widehat{e}^{SB}_{L[\beta]2}$ come from $\widehat{e}^{SB}_{L[\beta]2} = \beta\widehat{b}^{SB}_{L[\beta]2}$ and $\widehat{a}^{SB}_{L[\beta]2}$ comes from the fact that (16) is binding. ■

Proof of Proposition 5.

[1] $e^{SB}_{L[\beta]2} \leq \widehat{e}^{SB}_{L[\beta]2}$

We discuss 2 cases: $\widehat{e}^{SB}_{L[\beta]2} = \frac{\beta^2}{\gamma\sigma^2 + \widehat{\beta}^2 - 2\left(\widehat{\beta} - \beta\right)}$ and $\widehat{e}^{SB}_{L[\beta]2} = \beta$.

Case 1. $\widehat{e}^{SB}_{L[\beta]2} = \frac{\beta^2}{\gamma\sigma^2 + \widehat{\beta}^2 - 2\left(\widehat{\beta} - \beta\right)}$

$\gamma\sigma^2 + \widehat{\beta}^2 - 2\left(\widehat{\beta} - \beta\right) \geq \beta > 0$ and $\left(\gamma\sigma^2 + \widehat{\beta}^2 - 2\left(\widehat{\beta} - \beta\right)\right) - \left(\gamma\sigma^2 + \beta^2\right) = \left(\widehat{\beta} - \beta\right)\left(\widehat{\beta} + \beta - 2\right) < 0$, so $e^{SB}_{L[\beta]2} = \frac{\beta^2}{\gamma\sigma^2 + \beta^2} < \frac{\beta^2}{\gamma\sigma^2 + \widehat{\beta}^2 - 2\left(\widehat{\beta} - \beta\right)} = \widehat{e}^{SB}_{L[\beta]2}$.

Case 2. $\widehat{e}^{SB}_{L[\beta]2} = \beta$

$\widehat{e}^{SB}_{L[\beta]2} = \beta \geq \min\left\{\beta, \frac{\beta^2}{\gamma\sigma^2 + \beta^2}\right\} = e^{SB}_{L[\beta]2}$.

[2] $\widehat{e}^{SB}_{L[\beta]2} < e^{SB}_{[1]}$

We also discuss 2 cases: $\widehat{e}^{SB}_{L[\beta]2} = \frac{\beta^2}{\gamma\sigma^2+\widehat{\beta}^2-2\left(\widehat{\beta}-\beta\right)}$ and $\widehat{e}^{SB}_{L[\beta]2} = \beta$.

Case 1. $\widehat{e}^{SB}_{L[\beta]2} = \frac{\beta^2}{\gamma\sigma^2+\widehat{\beta}^2-2\left(\widehat{\beta}-\beta\right)}$

$e^{SB}_{[1]} - \widehat{e}^{SB}_{L[\beta]2} = \frac{1}{\gamma\sigma^2+1} - \frac{\beta^2}{\gamma\sigma^2+\widehat{\beta}^2-2\left(\widehat{\beta}-\beta\right)} = \frac{\gamma\sigma^2\left(1-\beta^2\right)+\widehat{\beta}^2-2\left(\widehat{\beta}-\beta\right)-\beta^2}{(1+\gamma\sigma^2)\left(\gamma\sigma^2+\widehat{\beta}^2-2\left(\widehat{\beta}-\beta\right)\right)}$.

From the fact that $\gamma\sigma^2 \geq \widehat{\beta}\left(1-\widehat{\beta}\right) + \left(\widehat{\beta}-\beta\right)$, we have $e^{SB}_{[1]} - \widehat{e}^{SB}_{L[\beta]2} \geq \frac{\left(-\widehat{\beta}^2+2\widehat{\beta}-\beta\right)\left(1-\beta^2\right)+\widehat{\beta}^2-2\left(\widehat{\beta}-\beta\right)-\beta^2}{(1+\gamma\sigma^2)\left(\gamma\sigma^2+\widehat{\beta}^2-2\left(\widehat{\beta}-\beta\right)\right)} = \frac{\beta\left((1-\beta)^2+\beta\left(1-\widehat{\beta}\right)^2\right)}{(1+\gamma\sigma^2)\left(\gamma\sigma^2+\widehat{\beta}^2-2\left(\widehat{\beta}-\beta\right)\right)} > 0$.

Case 2. $\widehat{e}^{SB}_{L[\beta]2} = \beta$

$e^{SB}_{[1]} - \widehat{e}^{SB}_{L[\beta]2} = \frac{1}{\gamma\sigma^2+1} - \beta = \frac{1-\beta-\beta\gamma\sigma^2}{\gamma\sigma^2+1}$.

From the fact that $\gamma\sigma^2 < \widehat{\beta}\left(1-\widehat{\beta}\right) + \left(\widehat{\beta}-\beta\right)$, we have $e^{SB}_{[1]} - \widehat{e}^{SB}_{L[\beta]2} > \frac{1-\beta-\beta\left(\widehat{\beta}\left(1-\widehat{\beta}\right)+\left(\widehat{\beta}-\beta\right)\right)}{\gamma\sigma^2+1} = \frac{(1-\beta)^2+\beta\left(1-\widehat{\beta}\right)^2}{\gamma\sigma^2+1} > 0$.

[3] $\pi^{SB}_{L[\beta]2} \leq \widehat{\pi}^{SB}_{L[\beta]2}$

From (11) and binding (16), we get

$$\widehat{\pi}^{SB}_{L[\beta]2} = \beta\widehat{b}^{SB}_{L[\beta]2} - \frac{1}{2}\left(\gamma\sigma^2 + \widehat{\beta}^2 - 2\left(\widehat{\beta}-\beta\right)\right)\left(\widehat{b}^{SB}_{L[\beta]2}\right)^2. \tag{18}$$

And from (11) and binding (14), we have $\pi^{SB}_{L[\beta]2} = \beta b^{SB}_{L[\beta]2} - \frac{1}{2}\left(\gamma\sigma^2+\beta^2\right)\left(b^{SB}_{L[\beta]2}\right)^2$.

Define $h(b) = \beta b - \frac{1}{2}\left(\gamma\sigma^2+\beta^2\right)b^2$ and $g(b) = \beta b - \frac{1}{2}\left(\gamma\sigma^2+\widehat{\beta}^2-2\left(\widehat{\beta}-\beta\right)\right)b^2$.

$g' > 0$ in $[0,1]$ as long as $\gamma\sigma^2 + \widehat{\beta}^2 - 2\left(\widehat{\beta}-\beta\right) < \beta$. When $\gamma\sigma^2 + \widehat{\beta}^2 - 2\left(\widehat{\beta}-\beta\right) \geq \beta > 0$, g is concave and $g' \geq 0$ in $\left[0, \frac{\beta}{\gamma\sigma^2+\widehat{\beta}^2-2\left(\widehat{\beta}-\beta\right)}\right]$.

h is always concave. And when $\gamma\sigma^2+\beta^2 < \beta$, $h' > 0$ in $[0,1]$; when $\gamma\sigma^2+\beta^2 \geq \beta$, $h' \geq 0$ in $\left[0, \frac{\beta}{\gamma\sigma^2+\beta^2}\right]$.

$g(b) - h(b) = -\frac{1}{2}\left(\widehat{\beta}^2 - \beta^2 - 2\left(\widehat{\beta}-\beta\right)\right)b^2 = -\frac{1}{2}\left(\widehat{\beta}-\beta\right)\left(\widehat{\beta}+\beta-2\right)b^2 > 0$ as long as $b \neq 0$.

We discuss 2 cases: $b_{L[\beta]2}^{SB} = \frac{\beta}{\gamma\sigma^2+\beta^2}$ and $b_{L[\beta]2}^{SB} = 1$.

Case 1. $b_{L[\beta]2}^{SB} = \frac{\beta}{\gamma\sigma^2+\beta^2}$

If $\widehat{b}_{L[\beta]2}^{SB} = \frac{\beta}{\gamma\sigma^2+\widehat{\beta}^2-2\left(\widehat{\beta}-\beta\right)}$, since $\gamma\sigma^2 + \widehat{\beta}^2 - 2\left(\widehat{\beta} - \beta\right) < \gamma\sigma^2 + \beta^2$, we have $\widehat{\pi}_{L[\beta]2}^{SB} = g\left(\widehat{b}_{L[\beta]2}^{SB}\right) > g\left(b_{L[\beta]2}^{SB}\right) > h\left(b_{L[\beta]2}^{SB}\right) = \pi_{L[\beta]2}^{SB}$.

If $\widehat{b}_{L[\beta]2}^{SB} = 1$, then $\gamma\sigma^2 + \widehat{\beta}^2 - 2\left(\widehat{\beta} - \beta\right) < \beta$, thus $g' > 0$ in $[0, 1]$. Hence $\widehat{\pi}_{L[\beta]2}^{SB} = g\left(1\right) \geq g\left(b_{L[\beta]2}^{SB}\right) > h\left(b_{L[\beta]2}^{SB}\right) = \pi_{L[\beta]2}^{SB}$.

Case 2. $b_{L[\beta]2}^{SB} = 1$

Based on the following proof of Proposition 7, we have $\widehat{b}_{L[\beta]2}^{SB} = 1$, then $\widehat{\pi}_{L[\beta]2}^{SB} = g\left(1\right) > h\left(1\right) = \pi_{L[\beta]2}^{SB}$.

[4] $\widehat{\pi}_{L[\beta]2}^{SB} < \pi_{[1]}^{SB}$

We know (18) and $\pi_{[1]}^{SB} = \frac{1}{2(\gamma\sigma^2+1)} = \frac{e_{[1]}^{SB}}{2}$. We discuss 2 cases: $\widehat{b}_{L[\beta]2}^{SB} = \frac{\beta^2}{\gamma\sigma^2+\widehat{\beta}^2-2\left(\widehat{\beta}-\beta\right)}$ and $\widehat{b}_{L[\beta]2}^{SB} = 1$.

Case 1. $\widehat{b}_{L[\beta]2}^{SB} = \frac{\beta^2}{\gamma\sigma^2+\widehat{\beta}^2-2\left(\widehat{\beta}-\beta\right)}$

We impose this into the expression of (18) and then get

$$\widehat{\pi}_{L[\beta]2}^{SB} = \frac{1}{2}\frac{\beta^2}{\gamma\sigma^2 + \widehat{\beta}^2 - 2\left(\widehat{\beta} - \beta\right)} = \frac{\widehat{e}_{L[\beta]2}^{SB}}{2},$$

so immediately from the result of [2], $\widehat{\pi}_{L[\beta]2}^{SB} < \pi_{[1]}^{SB}$.

Case 2. $\widehat{b}_{L[\beta]2}^{SB} = 1$

Now $\widehat{\pi}_{L[\beta]2}^{SB} = -\frac{1}{2}\gamma\sigma^2 - \frac{1}{2}\widehat{\beta}^2 + \widehat{\beta}$.

$\pi_{[1]}^{SB} - \widehat{\pi}_{L[\beta]2}^{SB} = \frac{1}{2}\left(\frac{1}{\gamma\sigma^2+1} + \gamma\sigma^2 + \widehat{\beta}^2 - 2\widehat{\beta}\right) = \frac{1}{2}\left(\frac{\left(\gamma\sigma^2\right)^2}{\gamma\sigma^2+1} + \left(1 - \widehat{\beta}\right)^2\right) > 0.$

■

Proof of Proposition 6.

(i)

Based on (18),

$$\begin{aligned}
\widetilde{\widehat{S}}^{SB}_{0L[\beta]2,P} &= \widehat{\pi}^{SB}_{L[\beta]2} + \frac{1}{\beta}\widehat{U}^{SB}_{0L[\beta]2} \\
&= \beta\widehat{b}^{SB}_{L[\beta]2} - \frac{1}{2}\left(\gamma\sigma^2 + \widehat{\beta}^2 - 2\left(\widehat{\beta} - \beta\right)\right)\left(\widehat{b}^{SB}_{L[\beta]2}\right)^2 \\
&\quad + \left(\left(\beta - \frac{\beta^2}{2}\right) - \left(\widehat{\beta} - \frac{\widehat{\beta}^2}{2}\right)\right)\left(\widehat{b}^{SB}_{L[\beta]2}\right)^2 \\
&= \beta\widehat{b}^{SB}_{L[\beta]2} - \frac{1}{2}\left(\gamma\sigma^2 + \beta^2\right)\left(\widehat{b}^{SB}_{L[\beta]2}\right)^2 \\
&= h\left(\widehat{b}^{SB}_{L[\beta]2}\right).
\end{aligned}$$

And from the proof of Proposition 5, $\pi^{SB}_{L[\beta]2} = \left(b^{SB}_{L[\beta]2}\right)$.

We then discuss 2 cases: $b^{SB}_{L[\beta]2} = \frac{\beta}{\gamma\sigma^2+\beta^2}$ and $b^{SB}_{L[\beta]2} = 1$.

Case 1. $b^{SB}_{L[\beta]2} = \frac{\beta}{\gamma\sigma^2+\beta^2}$

$b^{SB}_{L[\beta]2} = \frac{\beta}{\gamma\sigma^2+\beta^2}$ is at $h(\cdot)$'s global maximum, so whenever $\widehat{b}^{SB}_{L[\beta]2} = \frac{\beta^2}{\gamma\sigma^2+\widehat{\beta}^2-2\left(\widehat{\beta}-\beta\right)}$ or $\widehat{b}^{SB}_{L[\beta]2} = 1$, we have $h\left(b^{SB}_{L[\beta]2}\right) > h\left(\widehat{b}^{SB}_{L[\beta]2}\right)$, i.e., $\widetilde{\widehat{S}}^{SB}_{0L[\beta]2,P} < \pi^{SB}_{L[\beta]2} = \widetilde{S}^{SB}_{0L[\beta]2,P}$.

Case 2. $b^{SB}_{L[\beta]2} = 1$

From the following proof of Proposition 7, we know $\widehat{b}^{SB}_{L[\beta]2} = 1$, too. So $\widetilde{\widehat{S}}^{SB}_{0L[\beta]2,P} = \pi^{SB}_{L[\beta]2} = h(1)$.

Hence altogether $\widetilde{\widehat{S}}^{SB}_{0L[\beta]2,P} \leq \pi^{SB}_{L[\beta]2} = \widetilde{S}^{SB}_{0L[\beta]2,P}$.

(ii)

$$\begin{aligned}
\widehat{S}^{SB}_{0L[\beta]2} &= \widehat{\pi}^{SB}_{L[\beta]2} + \widehat{U}^{SB}_{0L[\beta]2} \\
&= \widetilde{\widehat{S}}^{SB}_{0L[\beta]2,P} - (1-\beta)\left(\left(\beta - \frac{\beta^2}{2}\right) - \left(\widehat{\beta} - \frac{\widehat{\beta}^2}{2}\right)\right)\left(\widehat{b}^{SB}_{L[\beta]2}\right)^2 \\
&= \widetilde{\widehat{S}}^{SB}_{0L[\beta]2,P} + \frac{1}{2}(1-\beta)\left((1-\beta)^2 - \left(1-\widehat{\beta}\right)^2\right)\left(\widehat{b}^{SB}_{L[\beta]2}\right)^2 \\
&> \widetilde{\widehat{S}}^{SB}_{0L[\beta]2,P}.
\end{aligned}$$

$(\beta, \gamma, \sigma^2) \in \Gamma \Rightarrow b^{SB}_{L[\beta]2} = 1 \Rightarrow \widehat{b}^{SB}_{L[\beta]2} = 1$ (from the proof of Proposition 7)$\Rightarrow \widetilde{\widehat{S}}^{SB}_{0L[\beta]2,P} = \pi^{SB}_{L[\beta]2} = h(1)$, so $\widehat{S}^{SB}_{0L[\beta]2} > \widetilde{\widehat{S}}^{SB}_{0L[\beta]2,P} = \pi^{SB}_{L[\beta]2} = S^{SB}_{0L[\beta]2}$. ■

Proof of Proposition 7. Under the definition of Γ and Ω, it suffices to prove $\forall \beta, \gamma, \sigma^2$, $b^{SB}_{L[\beta]2} = 1$ (i.e., $\widehat{\beta} = \beta$ and $(\beta, \gamma, \sigma^2) \in \Gamma$)$\Rightarrow \widehat{b}^{SB}_{L[\beta]2} = 1, \forall \widehat{\beta} > \beta$ (i.e., $\left(\beta, \widehat{\beta}, \gamma, \sigma^2\right) \in \Omega$).

We have $\left(\gamma\sigma^2 + \widehat{\beta}^2 - 2\left(\widehat{\beta} - \beta\right)\right) - (\gamma\sigma^2 + \beta^2) = \left(\widehat{\beta} - \beta\right)\left(\widehat{\beta} + \beta - 2\right) < 0$, so $b^{SB}_{L[\beta]2} = 1 \Rightarrow \beta \geq \gamma\sigma^2 + \beta^2 \Rightarrow \beta > \gamma\sigma^2 + \widehat{\beta}^2 - 2\left(\widehat{\beta} - \beta\right) \Rightarrow \widehat{b}^{SB}_{L[\beta]2} = 1$.

■

Figure 1.
Illustration of Quasi-hyperbolic Discounting

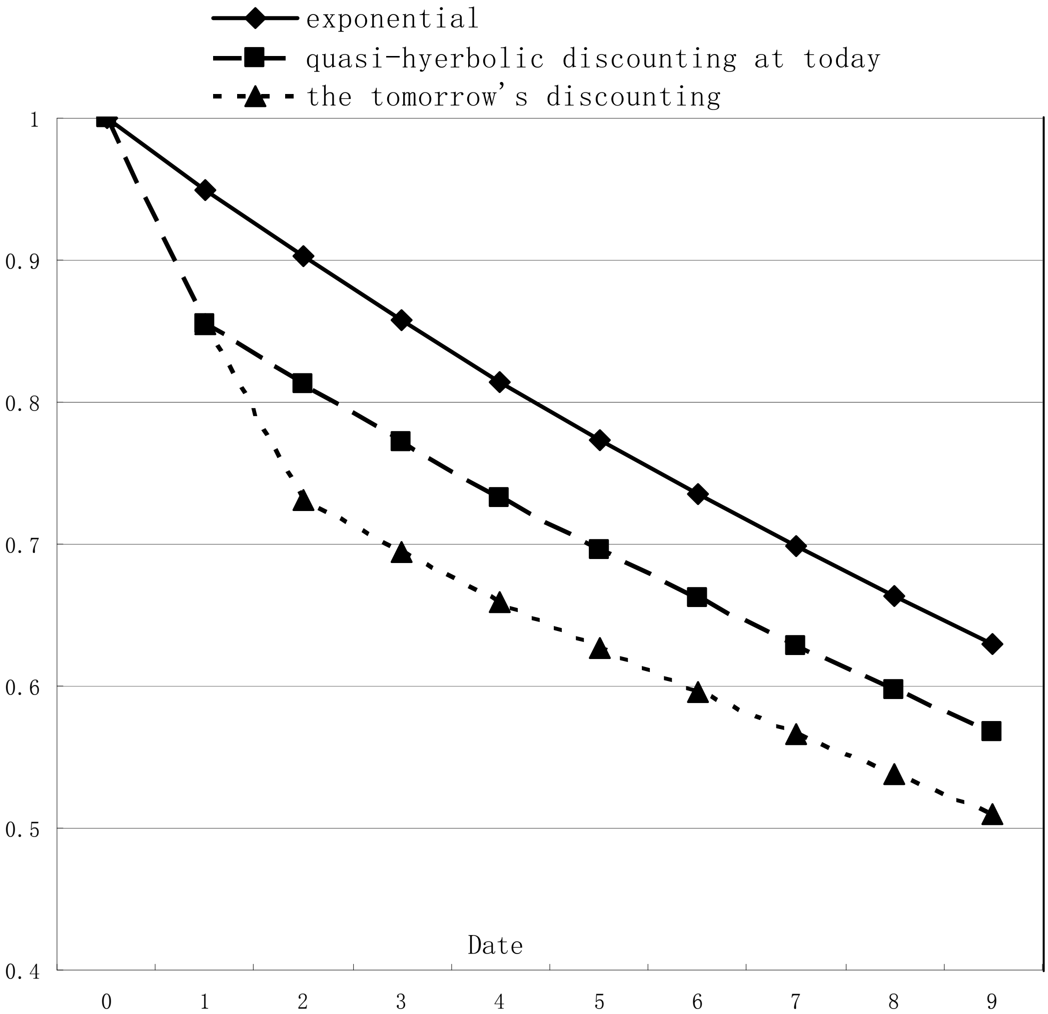


Figure 2 the Timing of the Benchmark (Short-run) Model

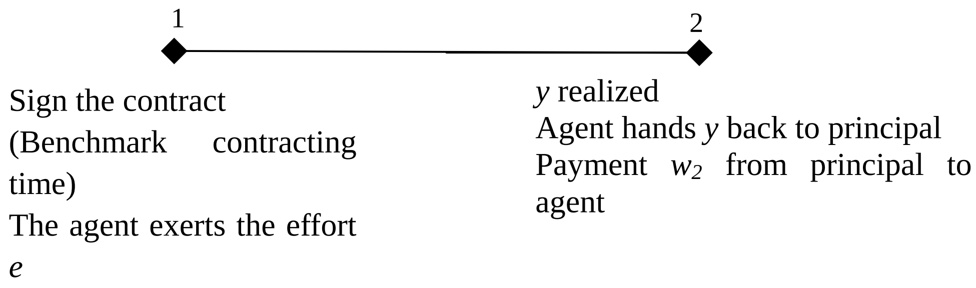


Figure 3 the Extended Timing of the Project

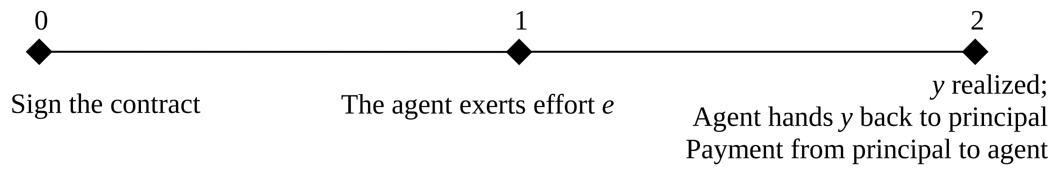


Figure 4 the Feature of the Contract for a Sophisticated Agent in the Long-run 2-sided Commitment Second-best

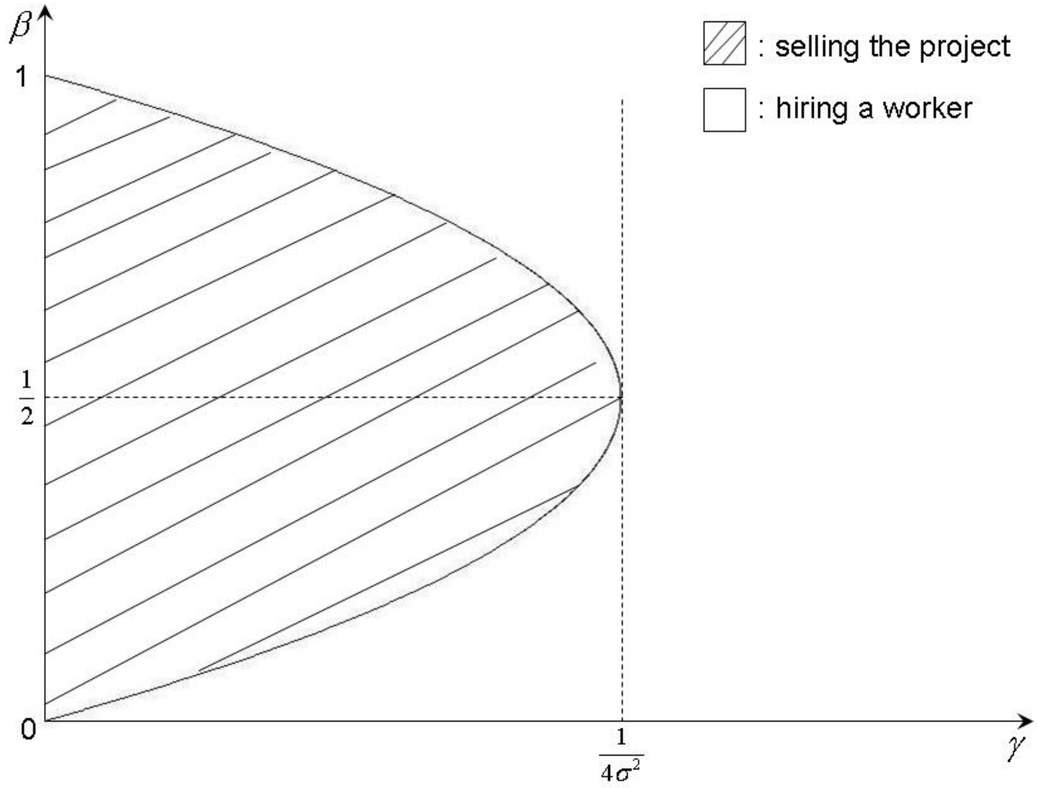
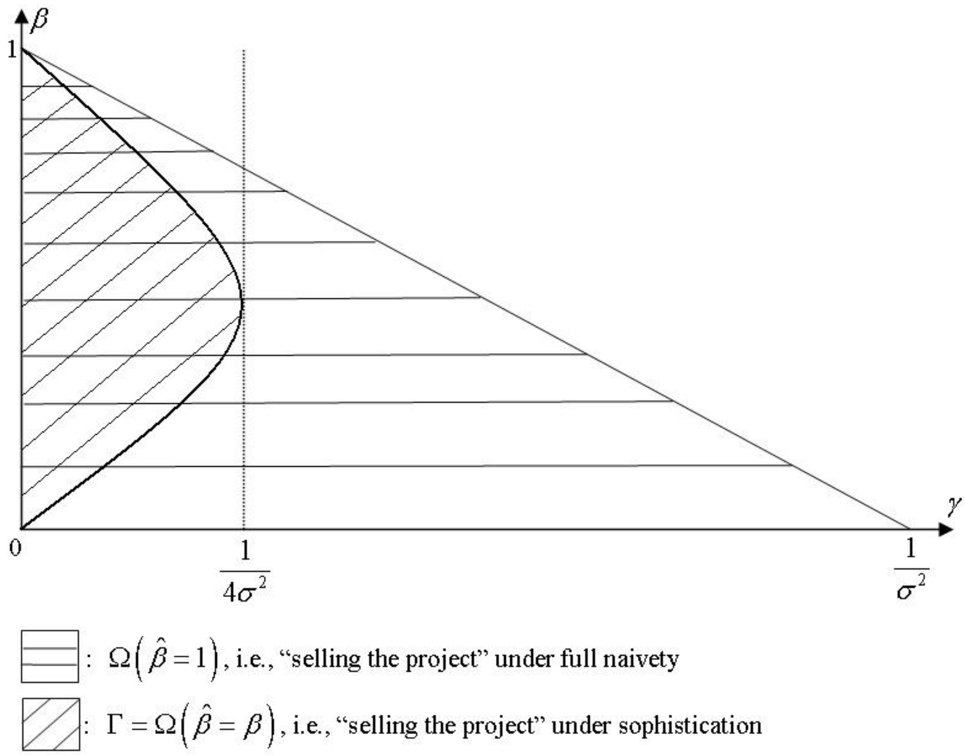


Figure 5 the Comparison between Sophistication and Full Naivety in the Long-run 2-sided Commitment Second-best



References

[1] Amador, Manuel, Iván Werning and George-Marios Angeletos, 2006, "Commitment vs. Flexibility", *Econometrica*, 74: 365-396.

[2] Bassi, Matteo, 2008, "Essays on Behavioral Public Economics", Ph.D. thesis, Toulouse School of Economics.

[3] Benabou, Roland and Jean Tirole, 2004, "Willpower and Personal Rules", *Journal of Political Economy*, 112: 848-886.

[4] Caillaud, B. and Jullien, B., 2000, "Modeling Time-inconsistent Preferences", *European Economic Review*, 44: 1116-1124.

[5] Carrillo, Juan D. and Thomas Mariotti, 2000, "Strategic Ignorance as A Self-Disciplining Device", *The Review of Economic Studies*, 67: 529-544.

[6] Chung, Shin-Ho, and Richard J. Herrnstein, 1961, "Relative and Absolute Strengths of Response as a Function of Frequency of Reinforcement", *Journal of the Experimental Analysis of Animal Behavior*, 4: 267-72.

[7] DellaVigna, S., and U. Malmendier, 2004, "Contract Design and Self-Control: Theory and Evidence", *Quarterly Journal of Economics*, 119: 353-402.

[8] Diamond, P. and B. Koszegi, 2003, "Quasi-hyperbolic Discounting and Retirement", *Journal of Public Economics*, 87: 1839-1872.

[9] Eliaz, Kfir and Ran Spiegler, 2006, "Contracting with Diversely Naive Agents", *The Review of Economic Studies*, 73: 689-714.

[10] Esteban, Susanna and Eiichi Miyagawa, 2006, "Temptation, self-control, and competitive nonlinear pricing", *Economics Letters*, 90: 348-355.

[11] Esteban, Susanna, Eiichi Miyagawa and Matthew Shum, 2007, "Nonlinear pricing with self-control preferences", *Journal of Economic Theory*, 135: 306-338.

[12] Gilpatric, Scott M., 2008, "Present-biased preferences, self-awareness and shirking", *Journal of Economic Behavior & Organization*, 67: 735-754.

[13] Gul, F. and Pesendorfer, W., 2001, "Temptation and Self Control", *Econometrica*, 69: 1403-1436.

[14] Harris, M. and A. Raviv, 1976, "Optimal Incentive Contracts with Imperfect Information", Working Paper #70-75-76, Graduate School of Industrial Administration, Carnegie-Mellon University (revised December 1977).

[15] Holmstrom, B., 1979, "Moral Hazard and Observability", *The Bell Journal of Economics*, 10: 74-91.

[16] Kylymnyuk, D., 2009, "Essays on Behavioral Economics and Contract Theory", Ph.D. thesis, Toulouse School of Economics.

[17] Laffont, J., and J. Tirole, 1993, *A Theory of Incentives in Procurement and Regulation*, MIT Press.

[18] Laibson, D., 1997, "Golden Eggs and Hyperbolic Discounting," *Quarterly Journal of Economics*, 112: 443-477.

[19] Luttmer, Erzo G. J. and Thomas Mariotti, 2007, "Efficiency and Equilibrium When Preferences Are Time-inconsistent", *Journal of Economic Theory*, 132: 493-506.

[20] O'Donoghue, Ted D. and Matthew Rabin, 2001, "Choice and Procrastination", *Quarterly Journal of Economics*, 116: 121-160.

[21] O'Donoghue, Ted and Matthew Rabin, 2005, "Optimal taxes for sin goods", *Swedish Economic Policy Review*, 12: 7-40.

[22] Phelps, E. S., and R. A. Pollak, 1968, "On Second-Best National Saving and Game-Equilibrium Growth", *Review of Economic Studies*, 35: 185-199.

[23] Samuelson, P., 1937, "A Note on Measurement of Utility", *The Review of Economic Studies*, 4: 155-161.

[24] Strotz, R., 1956, "Myopia and Inconsistency in Dynamic Utility Maximization", *The Review of Economic Studies*, 23: 165-180.