

SELF-REPUTATION AND PERCEPTION OF REPUTATION

Jung Hun Cho

CERGE-EI

Charles University
Center for Economic Research and Graduate Education
Academy of Sciences of the Czech Republic
Economics Institute

Working Paper Series
(ISSN 1211-3298)

343

**Self-Reputation and
Perception of Reputation**

Jung Hun Cho

CERGE-EI
Prague, November 2007

**ISBN 978-80-7343-142-6 (Univerzita Karlova. Centrum pro ekonomický výzkum
a doktorské studium)**
ISBN 978-80-7344-131-9 (Národohospodářský ústav AV ČR, v.v.i.)

Self-Reputation and Perception of Reputation*

Jung Hun Cho[†]

CERGE-EI[‡]

Abstract

This paper examines how the belief of decision maker regarding his ability to keep a resolution and his belief regarding what others think of him affect his actions. Higher self-reputation increases future payoff but higher perception of reputation can either increase or decrease it for an individual who has a strong ability to keep a resolution. However, both higher self-reputation and higher perception of reputation may not help increase future payoff for a decision maker who has a weak ability to resist temptation if he makes a resolution relatively easily in the second period. These results help to explain why some people ask for help or do not ask for help from friends to keep a resolution and why some people can or cannot sustain the resolution in the short run.

Abstrakt

Tato studie zkoumá, jak představy jedince o jeho schopnosti dodržet vlastní předsevzetí a jeho představy o tom, co si o této jeho schopnosti myslí okolí, ovlivňují jeho chování. Pro jedince se silnou schopností dodržet svá předsevzetí vede vyšší sebe-reputace ke zvýšeným budoucím výplatám, ale vyšší vnímání reputace může vést ke zvýšení i snížení budoucí výplaty. Vyšší sebe-reputace a vnímání reputace nemusí nutně zvýšit budoucí výplatu u jedince s nízkou schopností dodržovat svá předsevzetí, jestliže předsevzetí vzniká snadněji v druhém období. Výsledky napomáhají porozumět, proč někteří lidé ve snaze o dodržení předsevzetí žádají či nežadají o pomoc své okolí, a proč někteří lidé dodržují či nedodržují svá předsevzetí v krátkém období.

Keywords: Self-control; Self-reputation; Time inconsistency

JEL classification: D71; D82; D83

* I am grateful to Rajiv Sarin for many valuable comments and helpful discussions. I would like to thank Andreas Ortmann, Peter Katuscak and Avner Shaked for comments and suggestions.

[†] Correspondence address: CERGE-EI, P.O. Box 882, Politických vězňů 7, Prague, 111 21, Czech Republic. Tel.: (+420) 224 005 121; Fax: (+420) 224 227 143; E-mail: Junghun.Cho@cerge-ei.cz

[‡] CERGE-EI is a joint workplace of the Center for Economic Research and Graduate Education, Charles University in Prague, and the Economics Institute of Academy of Sciences of the Czech Republic.

1 Introduction

We often make promises to ourselves (resolutions) and try to keep them. Simple examples include deciding to quit smoking or starting a diet. However, many of us find it very hard to sustain resolutions, or even to make them in the first place because of the strong temptation to keep indulging in habit. Schelling (1960, 1978, 2006) introduces examples of commitments in the self-control problem. He uses commitments to mean becoming committed, bounded, or obligated to some course of action or to persist in inaction or some constraint on future action. The present paper starts from the observation that one possible way to stop a bad habit is to tell friends about the resolution. Indeed, on the website "Quit Smoking Stop" (<http://www.quit-smoking-stop.com/how-to-quit-smoking-tips.html>), visitors are urged to tell their family, friends and people they work with that they have left cigarettes.¹

Since exponential discounting ignores the tendency to obtain immediate rewards or to avoid immediate costs, most of the literature tries to explain the self-control problem by using hyperbolic discounting.² Hyperbolic discounting explains why people overvalue an immediate reward compared to a larger, later reward. In order to explain restrictions achieving long-term goals, either internal or external commitment devices can be examined. Laibson (1997) examines the external commitment device model of decision maker as a sequence of different selves. Similarly, O'Donoghue and Rabin (1999) examine the self-control problem by classifying a decision maker as either sophisticated or naive, and involvement in activities as either immediate costs or immediate rewards. Benabou and Tirole (2004) develop a model of internal commitment mechanisms by considering the self-reputation of decision maker. The present paper extends Benabou and

¹Similarly, the American Cancer Society (2005) mentions the importance of family support in quitting smoking.

²Strotz (1956) was the first economist to study dynamically inconsistent preferences. Pollak (1968), Peleg and Yaari (1973) and Goldman (1980) extended Strotz's work and examined an intrapersonal game among different temporal selves. See also Thaler and Shefrin (1981), Schelling (1978), Ainslie (1975, 2001), Laibson (1997), O'Donoghue and Rabin (1999), Brocas and Carrillo (2000).

Tirole's model by adding possibility that the decision maker can tell his friends about his commitment or not. This allows us to examine how both self-reputation and the belief regarding what the decision maker's friends think of him can change his choices.

Gul and Pesendorfer (2001, 2004) introduce an alternative approach without using hyperbolic discounting. They examine temptation rather than a preference change to explain self-control. Gul and Pesendorfer (2007) examine an infinite horizon consumption model to explain harmful addiction and obtain results consistent with the empirical evidence reported in Becker, Grossman and Murphy (1994) and Gruber and Koszegi (2001). They also examine the relationship between rehab and drug consumption and show that if there is a high cost to rehab, agents will never use it, and that there is therefore a cycle of increasing drug consumption.³ Fudenberg and Levine (2006) examine a dual-selves model in which there is a patient long-run self and a sequence of myopic short-run selves who have the same preferences but differ in how they regard the future.

Instead of allowing external constraints on the decision maker determine his ability to keep his resolution in the short run, I consider the possibility of telling friends about a resolution. I consider not only the effect of the decision maker's belief that he has a strong ability to keep a resolution but also the effect of the belief that his friends believe him to be a man of strong will. In addition, the subjective memory of his previous failures and the belief about the memory of his friends are examined to explain his choices. I ask: (a) assuming one has made a resolution, what does it take to sustain it in both the short run and the long run; (b) why is it that people keep making resolutions but (i) often succeed in keeping the resolution in the short run only and not in the long run, and (ii) sometimes cannot sustain it even in the short run.

In my model, a decision maker (DM) is uncertain about his ability to keep his resolution. I develop a two-period model. Each period is divided into two sub-periods. There are three choices in the first sub-period: continuing the habit, telling friends about

³Gruber and Koszegi (2001) start with a standard rational addiction model (Becker and Murphy, 1988) but incorporate time-inconsistent preferences.

one's resolution only after choosing not to continue the habit, and not telling friends after choosing not to continue the habit. In the second sub-period, the decision maker chooses whether to persevere or not. The first sub-period of each period can be thought of as the *Normal* period, and the second as the *Stress* period. The idea behind this distinction is that making resolutions, and telling about them is generally easier than keeping them. Figure 1 illustrates the structure of the decision problem.

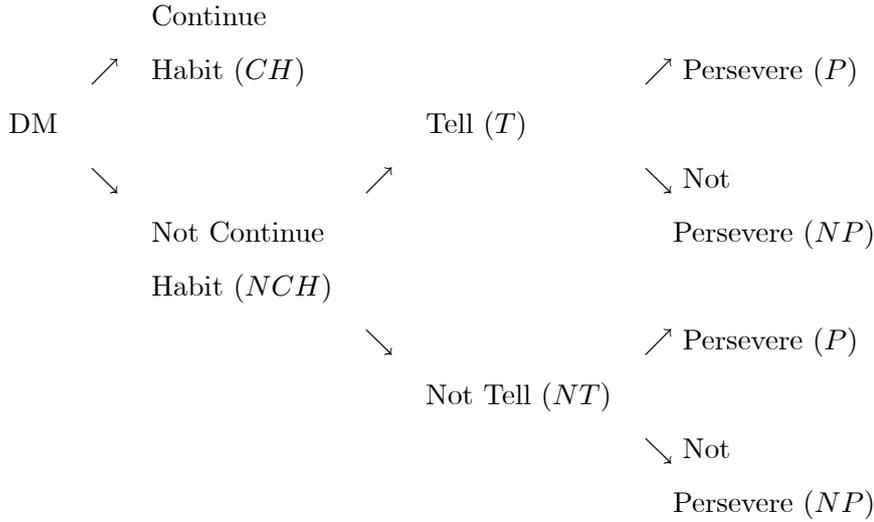


Figure 1- Decisions in each period

Figure 1 shows that the decision maker makes two ($\{\{CH\}, -, -\}$, $\{\{CH\}, -, -\}$), or four ($\{\{NCH\}, \{T/NT\}, \{P/NP\}\}$, $\{\{CH\}, -, -\}$), ($\{\{CH\}, -, -\}$, $\{\{NCH\}, \{T/NT\}, \{P/NP\}\}$), or six decisions ($\{\{NCH\}, \{T/NT\}, \{P/NP\}\}$, $\{\{NCH\}, \{T/NT\}, \{P/NP\}\}$) over the two periods. The decision maker can be one of two types: high type or low type. The high-type decision maker has a strong ability to sustain a resolution in the stress period while the low-type decision maker has a weak ability to keep a resolution in the stress period. If the decision maker chooses to continue his habit (*CH*) in the first period, it means that he has not made a resolution and is still uncertain about his type. Since the focus of the paper is to understand what it takes to sustain a resolution, I focus on the case in which the decision maker chooses not to continue his habit (*NCH*) in the first period. It is possible that the decision maker chooses not to continue habit

(*NCH*) in the second period but chooses not to persevere (*NP*) to the end. In this case, the decision maker is deemed to be a low type. The high-type decision maker is the individual who chooses not to continue his habit (*NCH*) and to persevere (*P*) in the second period. The difference between the types is captured by the difference in their discount rates in the stress period and is explained in detail in Sections 2 and 3.

Because of imperfect information regarding the type of decision maker, I examine two types of beliefs of the decision maker: self-reputation and perception of reputation. The decision maker believes that he is a high type with positive probability. The belief of the decision maker that he is a high type is termed *self-reputation*. The decision maker also believes that his friends believe him to be a high type with positive probability. The belief of the decision maker that his friends believe him to be a high type is called the *perception of reputation*. These two beliefs are updated by the choices the decision maker makes in the first period. In addition, if the decision maker chooses not to persevere in the first period, he will not remember his failure perfectly, i.e. with some probability he will forget that he failed to keep the resolution. As he does not remember his failure perfectly, he will interpret his previous choices in two different ways: either he succeeded in persevering or he never tried to test his willpower. Since friends may or may not listen to the resolution of the decision maker carefully, the decision maker believes that his friends will not remember what he promised perfectly, i.e. with some probability his friends will forget what he told them about his resolution.

Section 2 gives a detailed description of the model and its assumptions. Section 3 analyzes the model by using backward induction. The model yields interesting results, namely that:

(1) For the high-type decision maker, higher self-reputation increases his future payoff but higher perception of reputation can either increase or decrease it. Thus, he always chooses to persevere in the first period but he may or may not tell his friends about his resolution in the short run.

(2) For the low-type decision maker, if the ex-post payoff of continuing the habit in

the second period is sufficiently high, higher self-reputation increase his future payoff but higher perception of reputation decreases it. If the ex-post payoff of following temptation in the second period is sufficiently low, it is not always better to have either higher self-reputation or higher perception of reputation. This is because either higher self-reputation or higher perception of reputation can sometimes decrease the future payoff. In many cases, the low-type decision maker chooses not to tell his friends and not to persevere in the first period. But, there are cases in which he chooses not to tell his friends and not to persevere only if he considers his current period sufficiently important. If he considers his future sufficiently important, he sometimes chooses to tell his friends and not to persevere, or chooses to tell his friends and to persevere in the first period.

2 Model

I consider a two-period model in which each period is divided into two sub-periods. In the first sub-period of each period, when the decision maker chooses to continue his habit (CH), i.e. when he decides not to make a resolution, he can get the immediate benefit A_i in period $i \in \{1, 2\}$. In resolving to quit smoking, for example, the benefit of choosing CH measures the happiness smoking induces. In the first period, if the decision maker chooses CH , the decision of the first period is finished and the decision maker faces the choice in the next period.

If the decision maker chooses not to continue his habit (NCH), i.e. when he decides to make a resolution, he has to decide whether to tell his friends (T) or not to tell his friends (NT) about his resolution in the first sub-period of each period. It is assumed that there is no immediate cost from both the choice of NCH and the choice of T or NT in the first sub-period. In the second sub-period of each period, the decision maker has to choose to persevere (P) or not to persevere (NP). The benefit from not indulging in the habit is realized at the end of each period by following both the choice of T or NT and that of P or NP . In other words, after choosing NCH , the decision maker obtains

one of four possible delayed benefits in each period. When the decision maker chooses NCH , T (or NT) and P , he obtains the benefit $B_{i,T}$ (or $B_{i,NT}$) at the end of period $i \in \{1, 2\}$. When the decision maker chooses NCH , T (or NT) and NP , he obtains the benefit $b_{i,T}$ (or $b_{i,NT}$) at the end of period i . Since the decision maker resists temptation during the first sub-period, he suffers the immediate cost C_i in the second sub-period when he chooses P . But, if the decision maker chooses NP in the second sub-period, it is assumed that he does not suffer any cost.

Assumption 1.

$$B_{i,T} > B_{i,NT} > b_{i,NT} > b_{i,T} > A_i > 0 \text{ in period } i \in \{1, 2\}.$$

The benefit of choosing CH is an immediate payoff in period i , while the benefit of choosing NCH is delayed payoffs which are determined at the end point of period i . Regardless of telling friends or not, the benefit when the decision maker chooses to persevere ($B_{i,k}$) is greater than that when he chooses not to persevere ($b_{i,k}$) for $k \in \{T, NT\}$. Also, it is assumed that it is better for the decision maker to try to make a resolution than not, even though he cannot succeed in persevering in the second sub-period. Since the decision maker takes into account both self-reputation and the perception of reputation, the benefit of telling his friends about a resolution is greater than that of not telling his friends if he chooses P in the second sub-period. By using similar logic, the benefit of not telling his friends about his resolution is greater than that of telling his friends if he chooses NP in the second sub-period.

There is a standard discount rate δ between period 1 and period 2. In the model, I apply time-inconsistency preferences. The decision maker discounts the benefits that the choice of NCH would bring at rate γ . When the decision maker only considers his current period payoff, he chooses CH if $\frac{A_i}{\gamma}$ is greater than the expected payoff of choosing NCH . The discount rate γ is explained as the ability to keep a resolution in the normal period. Since this ability is well-known information to the decision maker (from his previous experiences), it is assumed for simplicity that the decision maker has perfect

information about γ . After the decision maker chooses NCH , he has an immediate cost and a delayed benefit in the second sub-period. I therefore apply the discount rate β in the second sub-period. The discount rate β is revealed only after the decision maker chooses NCH (i.e., he resists his impulses for some time before he confronts the next choice of P or NP). Since the decision maker will give up (i.e. chooses NP) whenever $\frac{C_i}{\beta} > B_{i,k} - b_{i,k}$ if he only considers his current period payoff, β is defined as the ability to keep the resolution in the stress period.

To examine both self-reputation and the perception of reputation, it is assumed that the decision maker is uncertain about both his ability to keep his resolution in the stress period and what his friends believe of his ability to keep it. The following assumptions are helpful to formalize these ideas.

Assumption 2.

The decision maker believes that he has a strong ability to keep his resolution in the stress period or is a high type in the stress period (β_H) with probability ρ_i in period $i \in \{1, 2\}$. With probability $1 - \rho_i$, the decision maker is a low type in the stress period ($\beta_L < \beta_H$).

In the first period, the decision maker has a prior belief ρ_1 that he is a high type in the stress period. The decision maker updates the belief regarding his own ability from ρ_1 to ρ_2 by the choices made during the first period. After choosing NCH , the choice P or NP plays an important role in determining the updated self-reputation. Even if the decision maker knows the value of β in the first period, he may not recall it perfectly in the second period.

Assumption 3.

With probability λ_s , the decision maker will remember that he chose not to persevere (NP) after choosing NCH in the first period. If the decision maker forgets his failure (with probability $1 - \lambda_s$), he will remember that he chose to persevere (P) after choosing

NCH with probability λ_{ss} and he will remember that he did not make the resolution (CH) with probability $1 - \lambda_{ss}$.

The decision maker will never forget his success (or P after NCH) but he may not remember his failure (NP after NCH) perfectly.⁴ If he does not remember his failure perfectly, he may have one of two wrong beliefs about his failure: He may believe that he chose to persevere after choosing not to continue his habit or he may believe that he never tried to test his willpower.

Assumption 4.

The decision maker believes that his friends believe him to be a high type (an individual who has discount factor β_H) with probability ξ_i in period i . With probability $1 - \xi_i$, the decision maker believes that his friends believe him to be a low-type individual.

In the first period, the decision maker has a prior belief ξ_1 . The decision maker updates his belief from ξ_1 to ξ_2 by the choices made during the first period. Especially, the choice T or NT plays an important role in determining the updated perception of reputation after choosing NCH .

Assumption 5.

With probability λ_f , the decision maker believes that his friends remember what he said during the first period.

From the point of view of friends, they may or may not care what the decision maker says about his resolution. Thus, the decision maker believes that his friends may remember what he promises with positive probability. In the analysis, this belief cannot change the choices of the decision maker in the second period.

⁴People tend to remember their successes more than their failures (Benabou and Tirole, 2004). See also Benabou and Tirole (2002).

3 The analysis

By using backward induction, the choices of the decision maker in each sub-period are examined. The critical values which guarantee changes in the choices are determined for each type of decision maker in each sub-period of each period.

3.1 Second period

Let us start by considering the second sub-period in the second period. For simplicity, it is assumed that the high-type decision maker in the stress period chooses to persevere (P), while the decision maker who has a weak ability to keep a resolution in the stress period chooses not to persevere (NP) in each period if he only considers his current period payoff.

Assumption 6.

$$\frac{C_i}{\beta_H} < B_{i,k} - b_{i,k} < \frac{C_i}{\beta_L}$$

for $k \in \{T, NT\}$.

Since the second period is the last period for the decision maker, he does not need to consider both self-reputation and the perception of reputation in this period. Regardless of telling friends or not, the high-type decision maker chooses to persevere (P) and the low-type decision maker chooses not to persevere (NP). This assumption implies that the decision maker chooses to persevere (P) with probability ρ_2 and chooses not to persevere (NP) with probability $1 - \rho_2$ where ρ_2 is the updated belief of the decision maker that he has a strong ability to keep his resolution in the stress period. I explain how this updated belief is determined later in the paper.

Now, let us determine the relationship between the updated belief of the decision maker that he is a high type in the stress period, ρ_2 , and the probability that the decision maker tells his friends about his resolution, $\phi(\rho_2, \xi_2)$ in the first sub-period of the second period. The decision maker chooses to tell his friends about his resolution (T)

if the net expected benefit from telling is greater than that from not telling. Given that the decision maker chooses not to continue his habit in the second period, he chooses to tell his friends about the resolution if

$$\rho_2(B_{2,T} - B_{2,NT}) + (1 - \rho_2)(b_{2,T} - b_{2,NT}) \geq 0.$$

If the updated belief of the decision maker that he is a high type in the stress period (ρ_2) is greater than the critical value (ρ_2^T), the decision maker chooses to tell his friends about the resolution, i.e. $\phi(\rho_2, \xi_2) = 1$ if $\rho_2 \geq \rho_2^T$ where

$$\rho_2^T = \frac{b_{2,NT} - b_{2,T}}{B_{2,T} - B_{2,NT} + b_{2,NT} - b_{2,T}}.$$

In order to examine the relationship between the updated belief of the decision maker that his friends believe him to be a high type, ξ_2 , and the probability that the decision maker tells his friends about the resolution, $\phi(\rho_2, \xi_2)$, I need the following assumption.

Assumption 7.

There exists ξ_2^* which satisfies

$$B_{i,T}^> > B_{i,NT}^> > b_{i,NT}^> >> b_{i,T}^> >> A_i^> > 0 \text{ if } \xi_2 \geq \xi_2^*$$

and

$$B_{i,T}^< >> B_{i,NT}^< > b_{i,NT}^< > b_{i,T}^< > A_i^< > > 0 \text{ if } \xi_2 < \xi_2^*$$

where

$$\begin{aligned} B_{i,T}^< &> B_{i,T}^> > B_{i,NT}^> > B_{i,NT}^< > > \\ b_{i,NT}^> &> b_{i,NT}^< > > b_{i,T}^< > > b_{i,T}^> > > A_i^< > > A_i^> . \end{aligned}$$

The assumption says that there is a critical value ξ_2^* which guarantees a change of the benefit in choices of the decision maker (happiness of the decision maker). If the

decision maker's belief that his friends believe him to be a high type is sufficiently high, the benefit of perseverance is not very big. It is because the decision maker thinks that his friends may not be surprised by his endurance or may perceive his endurance as being quite normal. But, his benefit from failing to persevere is very big when he does not tell his friends, compared to the case when tells his friends. By applying similar logic, if the belief of the decision maker that his friends believe him to be a high type is sufficiently low, his benefit of perseverance when he tells his friends is very big. It is because the decision maker thinks his friends may be surprised by his endurance, which in turn makes him happier. But, if he fails to persevere, telling his friends or not does not make big difference to the decision maker.

If the decision maker's belief that his friends believe him to be a high type is sufficiently high, since ρ_2^T can be increasing, the probability of telling his friends is decreased. This occurs because there is no big difference between telling and not telling to his friends when he perseveres, but it is a lot better for the decision maker not to tell his friends than to tell his friends if he fails to persevere. If the belief of the decision maker that his friends believe him to be a low type is sufficiently high, the probability of telling his friends is increased since it can lead to lower ρ_2^T . This is possible because telling his friends is a lot better than not telling his friends when he perseveres, but there is no big difference between telling and not telling if he fails to persevere.

In the first sub-period of the second period, the decision maker will choose not to continue his habit (*NCH*) if the net expected benefit of *NCH* is greater than the benefit of continuing. The benefit to the decision maker of continuing his habit is realized immediately and equal to A_2 . When the decision maker chooses not to continue the habit, then he chooses either to tell or not to tell his friends, and then chooses either to persevere or not to persevere. The expected benefit of not continuing his habit in the second period is

$$E_2(NCH) = \gamma[\phi(\rho_2, \xi_2)\{\rho_2(B_{2,T} - C_2) + (1 - \rho_2)b_{2,T}\} \\ + (1 - \phi(\rho_2, \xi_2))\{\rho_2(B_{2,NT} - C_2) + (1 - \rho_2)b_{2,NT}\}].$$

In Appendix A, it is shown that the benefit of not continuing the habit in the second period increases with the updated belief of the decision maker that he is a high type (ρ_2) if the benefit when the decision maker chooses to persevere (P) is sufficiently high. Formally,

$$\frac{\partial E_2(NCH)}{\partial \rho_2} > 0$$

if

$$B_{i,k} > C_i + b_{i,k}$$

for $k \in \{T, NT\}$. Because of this property, there exists ρ_2^* which satisfies $E_2(NCH) \geq A_2$ if $\rho_2 \geq \rho_2^*$ where

$$\begin{aligned} & \gamma[\phi(\rho_2^*, \xi_2)\{\rho_2^*(B_{2,T} - C_2) + (1 - \rho_2^*)b_{2,T}\} \\ & + (1 - \phi(\rho_2^*, \xi_2))\{\rho_2^*(B_{2,NT} - C_2) + (1 - \rho_2^*)b_{2,NT}\}] = A_2. \end{aligned}$$

The choices made by each type of decision maker in the first sub-period of the second period are summarized in the following proposition by comparing two critical values (ρ_2^T and ρ_2^*). If the ex-post payoff of choosing CH is sufficiently high (or low), i.e. if

$$\frac{A_2}{\gamma} \geq (\text{or } <) \frac{b_{2,NT}(B_{2,T} - C_2) - b_{2,T}(B_{2,NT} - C_2)}{B_{2,T} - B_{2,NT} + b_{2,NT} - b_{2,T}},$$

ρ_2^* is greater (or less) than or equal to ρ_2^T .

Proposition 1 1. *If the ex-post payoff of choosing CH ($\frac{A_2}{\gamma}$) is sufficiently high and $\rho_2 \geq \rho_2^*$, the decision maker chooses not to continue his habit (NCH) and chooses to tell his friends (T) in the second period.*

2. *If the ex-post payoff of choosing CH ($\frac{A_2}{\gamma}$) is sufficiently high and $\rho_2 < \rho_2^*$, the decision maker chooses to continue his habit (CH).*

This is the case in which the decision maker always tells his friends if he chooses not to continue his habit. The higher ex-post payoff of choosing CH makes the decision maker easily choose to continue the habit but if the updated belief that he has a strong

ability to keep his resolution is sufficiently high, he chooses not to continue his habit and chooses to tell his friends about his resolution.

Proposition 2 1. *If the ex-post payoff of choosing CH ($\frac{A_2}{\gamma}$) is sufficiently low and $\rho_2 \geq \rho_2^T$, the decision maker chooses not to continue his habit (NCH) and chooses to tell his friends (T) in the second period.*

2. *If the ex-post payoff of choosing CH ($\frac{A_2}{\gamma}$) is sufficiently low and $\rho_2^* < \rho_2 < \rho_2^T$, the decision maker chooses not to continue his habit (NCH) and chooses not to tell his friends (NT) in the second period.*

3. *If the ex-post payoff of choosing CH ($\frac{A_2}{\gamma}$) is sufficiently low and $\rho_2 < \rho_2^*$, the decision maker chooses to continue his habit (CH).*

This is the case in which the decision maker sometimes tells his friends after choosing not to continue his habit. Even if his ex-post payoff of choosing CH is sufficiently low, he chooses to continue the habit if his updated belief about his ability is sufficiently low. Again, the propositions are satisfied under the assumption that the high-type decision maker chooses P and the low-type decision maker chooses NP in the second sub-period of the second period.

Now, let us examine the value function of the decision maker as his net expected benefit from the second period decisions. Since the choice of each type of decision maker in the second period is assumed differently, I need to separate the value function for each type. The value function for the decision maker who has a strong ability to keep a resolution in the stress period (i.e., who is a high type) is

$$V_2^H(\rho_2, \xi_2) = P(\rho_2, \xi_2)[\phi(\rho_2, \xi_2)(B_{2,T} - C_2) + (1 - \phi(\rho_2, \xi_2))(B_{2,NT} - C_2)] + (1 - P(\rho_2, \xi_2))A_2$$

where $P(\rho_2, \xi_2)$ is the probability that the decision maker chooses NCH in the second period which is determined by the critical value ρ_2^* . Similarly, the value function of the

decision maker who has a weak ability to keep a resolution in the stress period (i.e., who is a low type) is

$$V_2^L(\rho_2, \xi_2) = P(\rho_2, \xi_2)[\phi(\rho_2, \xi_2)b_{2,T} + (1 - \phi(\rho_2, \xi_2))b_{2,NT}] + (1 - P(\rho_2, \xi_2))A_2.$$

In order to examine the properties of the value function, let us first consider the case in which the ex-post payoff of choosing CH is sufficiently high. If the updated belief of the decision maker that he has a strong ability to keep his resolution is greater than the critical value ρ_2^* , the value function for each type is

$$V_2^H(\rho_2, \xi_2) = B_{2,T} - C_2$$

and

$$V_2^L(\rho_2, \xi_2) = b_{2,T}.$$

But, if ρ_2 is less than ρ_2^* , the value function for each type is

$$V_2^H(\rho_2, \xi_2) = A_2 = V_2^L(\rho_2, \xi_2).$$

Regardless of the type of decision maker, the value function increases with ρ_2 from below ρ_2^* to above ρ_2^* . From Assumption 7, it is possible to examine the properties of the value function with ξ_2 for each type. Regardless of the type, the value function decreases with ξ_2 from below ξ_2^* to above ξ_2^* .

Next is the case in which the ex-post payoff of choosing CH is sufficiently low. In this case, there are three possible value functions for each type. If the updated belief of the decision maker that he has a strong ability to keep his resolution is greater than the critical value of telling his friends, ρ_2^T , then the value function for each type is

$$V_2^H(\rho_2, \xi_2) = B_{2,T} - C_2$$

and

$$V_2^L(\rho_2, \xi_2) = b_{2,T}.$$

If ρ_2 is between ρ_2^* and ρ_2^T , the value function for each type is

$$V_2^H(\rho_2, \xi_2) = B_{2,NT} - C_2$$

and

$$V_2^L(\rho_2, \xi_2) = b_{2,NT}.$$

Finally, if ρ_2 is less than ρ_2^* , the value function for each type is

$$V_2^H(\rho_2, \xi_2) = A_2 = V_2^L(\rho_2, \xi_2).$$

In this case, the value function for the high type increases with ρ_2 (from $\rho_2 < \rho_2^*$ to $\rho_2^* < \rho_2 < \rho_2^T$ and from $\rho_2^* < \rho_2 < \rho_2^T$ to $\rho_2 > \rho_2^T$). However, the value function for the low type increases with ρ_2 (from $\rho_2 < \rho_2^*$ to $\rho_2^* < \rho_2 < \rho_2^T$) but decreases with ρ_2 (from $\rho_2^* < \rho_2 < \rho_2^T$ to $\rho_2 > \rho_2^T$). Higher self-reputation is always better for the high-type decision maker, but it is not always better for the low-type decision maker. It is a shock for the decision maker who believes himself to be a high type in the normal period to realize that he is not a high type in the stress period. However, this shock is weaker if the belief of the decision maker that he is a high type in the stress period is lower. This is because the low-type decision maker in the stress period chooses not to persevere if he only considers his current payoff. Thus, the low-type decision maker in the stress period feels comfortable choosing *NCH* if he believes that he is a low type. From Assumption 7, it is also possible to examine the properties of the value function with ξ_2 for each type. Under the condition that ρ_2 is less than ρ_2^* or ρ_2 is greater than ρ_2^T , the value function decreases with the updated perception of reputation from below ξ_2^* to above ξ_2^* . However, it increases with the updated perception of reputation from below ξ_2^* to above ξ_2^* if ρ_2 lies between ρ_2^* and ρ_2^T . The relationship between the value function and the updated beliefs are explained later in the paper when the choices of each type of the decision maker are examined.

3.2 First period

The updated beliefs of the decision maker are determined by his choices in the first period, the probability of remembering his failure, and his belief about the memory of his friends. When the decision maker chooses CH in the first period, then he obtains neither the opportunity to tell his friends nor the opportunity to test his strength of will in the stress time. As a result, both ρ_1 and ξ_1 remain unchanged, i.e. $\rho_1 = \rho_2$ and $\xi_1 = \xi_2$. If he chooses NCH and NP in the first period, but does not remember his failure and remembers that he never tried to test his willpower, then ρ_1 is not changed. When he chooses NCH , if he chooses not to tell his friends, or if he chooses to tell his friends but believes that his friends will forget what he promised, ξ_1 is not changed.

The self-reputation increases from ρ_1 to ρ_2^+ if the decision maker chooses to persevere (P) in the first period regardless of telling his friends or not after choosing NCH . Even if he chooses NP in the first period, the false memory that he chose P will lead to increasing self-reputation ρ_2^+ . Only if the decision maker remembers that he chose NP in the first period, the self-reputation will decrease from ρ_1 to ρ_2^- .

The perception of reputation is changed if the decision maker chooses to tell his friends about his resolution and if he believes that his friends will remember what he told them in the first period. It increases from ξ_1 to ξ_2^+ if the decision maker tells his friends and then chooses to persevere after choosing NCH , but decreases to ξ_2^- if he tells his friends and then chooses not to persevere after choosing NCH in the first period.

Now, let us calculate the total payoff of each type of decision maker and then examine choices in the first period. The total payoff of the decision maker who is j type for $j \in \{\text{High } (H), \text{Low } (L)\}$ when he chooses T and P in the first period is

$$B_{1,T} - \frac{C_1}{\beta_H} + \delta[\lambda_f V_2^j(\rho_2^+, \xi_2^+) + (1 - \lambda_f) V_2^j(\rho_2^+, \xi_1)].$$

Because of Assumption 6, the high-type decision maker chooses to persevere while the low-type decision maker chooses not to persevere in the first period if he only considers

his first period. In the first period, since he momentarily knows his type when he decides whether to persevere or not, his expected value function is determined as $\lambda_f V_2^j(\rho_2^+, \xi_2^+) + (1 - \lambda_f) V_2^j(\rho_2^+, \xi_1)$.

When the decision maker chooses NT and P in the first period, the total payoff of each type of decision maker is

$$B_{1,NT} - \frac{C_1}{\beta_H} + \delta V_2^j(\rho_2^+, \xi_1).$$

Similarly, the total payoffs of choosing T and NP and that of choosing NT and NP are

$$\begin{aligned} & b_{1,T} + \delta[\lambda_s\{\lambda_f V_2^j(\rho_2^-, \xi_2^-) + (1 - \lambda_f) V_2^j(\rho_2^-, \xi_1)\} \\ & + (1 - \lambda_s)[\lambda_{ss}\{\lambda_f V_2^j(\rho_2^+, \xi_2^+) + (1 - \lambda_f) V_2^j(\rho_2^+, \xi_1)\} \\ & + (1 - \lambda_{ss}) V_2^j(\rho_1, \xi_1)] \end{aligned}$$

and

$$\begin{aligned} & b_{1,NT} + \delta[\lambda_s V_2^j(\rho_2^-, \xi_1) \\ & + (1 - \lambda_s)\{\lambda_{ss} V_2^j(\rho_2^+, \xi_1) + (1 - \lambda_{ss}) V_2^j(\rho_1, \xi_1)\}] \end{aligned}$$

respectively.

The choice CH or NCH in the first sub-period of the first period is determined regardless of the type of decision maker because he knows his type only after he chooses NCH in the model. Thus, regardless of the type of decision maker, the total payoff of the decision maker when he chooses NCH is

$$\begin{aligned} & \gamma[\phi(\rho_1, \xi_1)\{Q(\rho_1, \xi_1)(B_{1,T} - C_1) + (1 - Q(\rho_1, \xi_1))b_{1,T}\} \\ & + (1 - \phi(\rho_1, \xi_1))\{Q(\rho_1, \xi_1)(B_{1,NT} - C_1) + (1 - Q(\rho_1, \xi_1))b_{1,NT}\}] \\ & + \delta[\rho_1 V_2^H(\rho_2, \xi_2) + (1 - \rho_1) V_2^L(\rho_2, \xi_2)], \end{aligned}$$

where $\phi(\rho_1, \xi_1)$ is the probability that the decision maker tells his friends in the first period and $Q(\rho_1, \xi_1)$ is the probability that the decision maker chooses NP in the first

period. These two probabilities are determined by the choices of each type of decision maker in equilibrium. The equilibrium conditions are explained in the next section. Similarly, the total payoff of the decision maker when he chooses CH in the first period is

$$A_1 + \delta[\rho_1 V_2^H(\rho_2, \xi_2) + (1 - \rho_1) V_2^L(\rho_2, \xi_2)].$$

Thus, each type of decision maker chooses NCH if $\rho_1 \geq \rho_1^*$ where ρ_1^* satisfies

$$\begin{aligned} & \gamma[\phi(\rho_1^*, \xi_1)\{Q(\rho_1^*, \xi_1)(B_{1,T} - C_1) + (1 - Q(\rho_1^*, \xi_1))b_{1,T}\} \\ & + (1 - \phi(\rho_1^*, \xi_2))\{Q(\rho_1^*, \xi_1)(B_{1,NT} - C_1) + (1 - Q(\rho_1^*, \xi_1))b_{1,NT}\}] = A_1. \end{aligned}$$

If the decision maker chooses to continue his habit, he confronts the choices in the second period explained in the previous section. Let us consider the case in which the decision maker chooses not to continue his habit in the first period, i.e. $\rho_1 \geq \rho_1^*$.

3.2.1 First period decisions of the high type

The choices of each type of decision maker in the first period are separated into two different cases: $\rho_2^T \leq \rho_2^*$ or $\rho_2^T > \rho_2^*$ (the ex-post payoff of choosing CH in the second period is sufficiently high or low). In each case, the choices of the decision maker in the first period are examined by prior beliefs, ρ_1 and ξ_1 . When the decision maker who has a strong ability to sustain a resolution in the stress period chooses to persevere (P) in the second sub-period of the first period, he can increase both the payoff in the first period and the value function. The choice of T or NT in the first sub-period of the first period is explained in proposition 3 and proposition 4.

Proposition 3 *The high-type decision maker in the stress period chooses to tell his friends about his resolution (T) and chooses to persevere (P) after choosing not to continue his habit (NCH) in the first period if $\xi_2^+ > \xi_1 > \xi_2^- > \xi_2^*$ or $\xi_2^* > \xi_2^+ > \xi_1 > \xi_2^-$ are satisfied(See Appendix B).*

Compared to the critical value ξ_2^* , if the prior belief that his friends believe him to be a high type is either very high or very low (or if the choice of P or NP after choosing T can slightly change the perception of reputation), the high-type decision maker chooses to tell his friends and to persevere in the first period. In these cases, the value function is not affected by the change in the perception of reputation. But, to tell his friends and to persevere lead to the highest payoff in the first period; T and P is the dominant strategy.

But, if the perception of reputation in failure falls below the critical value when the prior belief is above the critical value, or if the perception of reputation in success increases above the critical value when the prior belief is below it (i.e., if $\xi_2^+ > \xi_1 > \xi_2^* > \xi_2^-$ or $\xi_2^+ > \xi_2^* > \xi_1 > \xi_2^-$), the high-type decision maker sometimes tells his friends about his resolution.

Proposition 4 *Under the condition $\xi_2^+ > \xi_1 > \xi_2^* > \xi_2^-$ or $\xi_2^+ > \xi_2^* > \xi_1 > \xi_2^-$, the high-type decision maker in the stress period does not always choose to tell his friends about his resolution (T) and chooses to persevere (P) after choosing not to continue his habit (NCH) in the first period. He sometimes chooses not to tell his friends (NT) and to persevere (P) in the first period (See Appendix C).*

The critical value of the discount rate which guarantees a change in the choice of the decision maker from T and P to NT and P is determined. The high-type decision maker who considers his current period sufficiently important chooses T and P because this is the case in which the decrease in the value function from the updated perception of reputation is greater than the increase in the value function from the updated self-reputation. So, in this case, if he considers his future sufficiently important, he chooses NT and P in the first period.

In many cases (e.g. $\rho_2^+ > \rho_1 > \rho_2^* > \rho_2^- > \rho_2^T$ and $\xi_2^+ > \xi_1 > \xi_2^* > \xi_2^-$, or $\rho_2^T > \rho_2^+ > \rho_1 > \rho_2^*$ and $\xi_2^+ > \xi_2^* > \xi_1 > \xi_2^-$, etc.), the high-type decision maker always chooses T

and P in the first period. Let us consider the case in which $\rho_2^+ > \rho_1 > \rho_2^- > \rho_2^* > \rho_2^T$ and $\xi_2^+ > \xi_2^* > \xi_1 > \xi_2^-$ are satisfied. The high-type decision maker chooses T and P rather than NT and P if

$$\delta \leq \frac{B_{1,T} - \frac{C_1}{\beta_H} - b_{1,T}}{\lambda_f \{1 - \lambda_{ss}(1 - \lambda_s)\} (B_{2,T}^< - B_{2,T}^>)} = \delta^*,$$

but he chooses NT and P if $\delta > \delta^*$. This is the case in which the change in the value function from the updated perception of reputation is greater than that in the value function from the updated self-reputation. If the high-type decision maker considers his first period sufficiently important, he chooses T and P . But, if telling his friends about the resolution is not helpful for the high-type decision maker, he chooses NT and P in the first period. Also, the probability of choosing T and P increases with $\lambda_{ss}(1 - \lambda_s)$ and decreases with λ_f , i.e. the high-type decision maker chooses T and P more easily if his belief that he succeeds in persevering even if he fails to persevere increases, or if his belief that his friends will remember what he promises decreases. I will not explain other cases because the analysis is the same, but only the critical value which guarantees the change of the choice of the decision maker is different. Finally, the high-type decision maker never chooses NP in the first period because this choice can decrease both his current payoff and his future payoff.

3.2.2 First period decisions of the low type

The low-type decision maker chooses not to persevere if he only considers his current period payoff. The value function of the low-type decision maker increases or decreases with both updated beliefs of the decision maker. Let us first examine the case in which the ex-post payoff of choosing CH is sufficiently high, i.e. ρ_2^* is greater than or equal to ρ_2^T . In this case, the value function of the low-type decision maker increases with the updated self-reputation.

If the updated belief that he is a high type when he chooses not to persevere is greater than the critical value ρ_2^* , and the updated perception of reputation is either high

enough or low enough (i.e. if $\rho_2^+ > \rho_1 > \rho_2^- > \rho_2^* > \rho_2^T$ and either $\xi_2^+ > \xi_1 > \xi_2^- > \xi_2^*$ or $\xi_2^* > \xi_2^+ > \xi_1 > \xi_2^-$), NT and NP is the dominant strategy. In these cases, the value function is not changed by the perception of reputation. Since the low-type decision maker maximizes his first period payoff when he chooses NT and NP and his expected payoff slightly falls in this case (because of $\rho_2^- > \rho_2^*$), NT and NP in the first period is the dominant strategy.

Let us examine the case in which the low-type decision maker believes that $\rho_2^+ > \rho_1 > \rho_2^- > \rho_2^* > \rho_2^T$ and $\xi_2^+ > \xi_1 > \xi_2^* > \xi_2^-$ are satisfied. The total payoff of the low-type decision maker when he chooses T (or NT) and P is

$$B_{1,T} - \frac{C_1}{\beta_L} + \delta b_{2,T}^>$$

$$(\text{or } B_{1,NT} - \frac{C_1}{\beta_L} + \delta b_{2,T}^>).$$

Similarly, the total payoff of choosing T (or NT) and NP is

$$b_{1,T} + \delta[\lambda_s\{\lambda_f b_{2,T}^< + (1 - \lambda_f)b_{2,T}^>\} + (1 - \lambda_s)b_{2,T}^>]$$

$$(\text{or } b_{1,NT} + \delta b_{2,T}^>).$$

Regardless of δ and either the choice T or NT , the low-type decision maker chooses NP in the first period. But if

$$\delta \leq \frac{b_{1,NT} - b_{1,T}}{\lambda_s \lambda_f (b_{2,T}^< - b_{2,T}^>)},$$

the low-type decision maker chooses NT and NP rather than T and NP . In this case, the value function increases with the updated perception of reputation when the decision maker chooses T and NP . So, if he considers his future payoff sufficiently important, he chooses T and NP . If the decision maker's belief that his friends remember what he promised falls because telling his friends is not helpful to him, the probability of choosing NT and NP increases.

In the case where $\rho_2^+ > \rho_1 > \rho_2^- > \rho_2^* > \rho_2^T$ and $\xi_2^+ > \xi_2^* > \xi_1 > \xi_2^-$ are satisfied, the low-type decision maker never chooses T and P because it can decrease his current payoff

and his future payoff (since the decrease in the value function from the updated perception of reputation is greater than the increase in the value function from the updated self-reputation). But, the value function from the updated perception of reputation is not changed if he chooses NP after choosing either T or NT . Since he can maximize his first period payoff by choosing NT and NP and his expected payoff slightly falls in this case, he always chooses NT and NP in the first period.

If self-reputation falls below the critical value and this fall can decrease the value function, the decision maker sometimes chooses T and P , especially if he considers his future sufficiently important. Let us consider the case in which the decision maker believes that either $\rho_2^+ > \rho_1 > \rho_2^* > \rho_2^- > \rho_2^T$ or $\rho_2^+ > \rho_1 > \rho_2^* > \rho_2^T > \rho_2^-$, and either $\xi_2^+ > \xi_1 > \xi_2^- > \xi_2^*$ or $\xi_2^* > \xi_2^+ > \xi_1 > \xi_2^-$ are satisfied. Since both ρ_2^+ and ρ_1 are greater than the critical value and ρ_2^- is less than it, the value function for each ρ is

$$V_2^L(\rho_2^+, \xi_2) = b_{2,T} = V_2^L(\rho_1, \xi_2)$$

and

$$V_2^L(\rho_2^-, \xi_2) = A_2.$$

The total payoff of the low-type decision maker when he chooses T (or NT) and P is

$$\begin{aligned} & B_{1,T} - \frac{C_1}{\beta_L} + \delta b_{2,T}^> \\ & (\text{or } B_{1,NT} - \frac{C_1}{\beta_L} + \delta b_{2,T}^>) \end{aligned}$$

if $\xi_2^+ > \xi_1 > \xi_2^- > \xi_2^*$.

Similarly, the total payoff of choosing T (or NT) and NP is

$$\begin{aligned} & b_{1,T} + \delta[\lambda_s A_2^> + (1 - \lambda_s) b_{2,T}^>] \\ & (\text{or } b_{1,NT} + \delta[\lambda_s A_2^> + (1 - \lambda_s) b_{2,T}^>]) \end{aligned}$$

if $\xi_2^+ > \xi_1 > \xi_2^- > \xi_2^*$.

Proposition 5 *Under the conditions that $\xi_2^+ > \xi_1 > \xi_2^- > \xi_2^*$ and either $\rho_2^+ > \rho_1 > \rho_2^* > \rho_2^- > \rho_2^T$ or $\rho_2^+ > \rho_1 > \rho_2^* > \rho_2^T > \rho_2^-$ are satisfied, there is*

$$\delta^{**} = \frac{b_{1,NT} - B_{1,T} + \frac{C_1}{\beta_L}}{\lambda_s(b_{2,T}^> - A_2^>)}$$

*such that if $\delta > \delta^{**}$, the low-type decision maker chooses T and P and if $\delta < \delta^{**}$, he chooses NT and NP in the first period (See Appendix D).*

There is a critical discount value which guarantees the change of choices from T and P to NT and NP . Since the low-type decision maker can increase the value function by choosing T and P in this case, if he considers his future sufficiently important, he never chooses NT and NP . If the probability of remembering his failure increases, the probability of choosing T and P increases. Similarly, the critical value of δ ,

$$\frac{b_{1,NT} - B_{1,T} + \frac{C_1}{\beta_L}}{\lambda_s(b_{2,T}^< - A_2^<)},$$

is also determined if the low-type decision maker believes $\xi_2^* > \xi_2^+ > \xi_1 > \xi_2^-$ and either $\rho_2^+ > \rho_1 > \rho_2^* > \rho_2^- > \rho_2^T$ or $\rho_2^+ > \rho_1 > \rho_2^* > \rho_2^T > \rho_2^-$ are satisfied. If the discount rate is greater than this critical value, the low-type decision maker chooses T and P ; otherwise he chooses NT and NP in the first period. By using the same method, it is shown why the low-type decision maker chooses NT and NP , T and NP , or T and P in the first period under the condition that ρ_2^* is greater than or equal to ρ_2^T .⁵ The low-type decision maker never chooses NT and P in the first period because he can never increase both his current payoff and his future payoff by doing so.

Next is the case in which the ex-post payoff of choosing CH in the second period is sufficiently low, i.e. ρ_2^* is less than ρ_2^T . In this case, the value function for the low type increases with ρ_2 (from $\rho_2 < \rho_2^*$ to $\rho_2^* < \rho_2 < \rho_2^T$) but decreases with ρ_2

⁵Similarly, the cases where $\rho_2^+ > \rho_2^* > \rho_1 > \rho_2^T$, $\rho_2^* > \rho_2^+ > \rho_1 > \rho_2^T$, $\rho_2^+ > \rho_2^* > \rho_2^T > \rho_1$, $\rho_2^* > \rho_2^+ > \rho_2^T > \rho_1$, or $\rho_2^* > \rho_2^T > \rho_2^+ > \rho_1$ is satisfied under 4 different cases ($\xi_2^+ > \xi_1 > \xi_2^- > \xi_2^*$, $\xi_2^+ > \xi_1 > \xi_2^* > \xi_2^-$, $\xi_2^+ > \xi_2^* > \xi_1 > \xi_2^-$ and $\xi_2^* > \xi_2^+ > \xi_1 > \xi_2^-$) are examined to determine the choices of the decision maker and the critical discount rate.

(from $\rho_2^* < \rho_2 < \rho_2^T$ to $\rho_2 > \rho_2^T$), and high perception of reputation is not always better for the low-type decision maker. If the prior self-reputation is greater than ρ_2^T , the results are similar to the previous cases.⁶ So, let us consider the case in which the prior self-reputation lies between ρ_2^* and ρ_2^T . There are four different cases in ρ - $\rho_2^+ > \rho_2^T > \rho_1 > \rho_2^* > \rho_2^-$, $\rho_2^+ > \rho_2^T > \rho_1 > \rho_2^- > \rho_2^*$, $\rho_2^T > \rho_2^+ > \rho_1 > \rho_2^* > \rho_2^-$ and $\rho_2^T > \rho_2^+ > \rho_1 > \rho_2^- > \rho_2^*$ under four different cases in ξ - $\xi_2^+ > \xi_1 > \xi_2^- > \xi_2^*$, $\xi_2^+ > \xi_1 > \xi_2^* > \xi_2^-$, $\xi_2^+ > \xi_2^* > \xi_1 > \xi_2^-$ and $\xi_2^* > \xi_2^+ > \xi_1 > \xi_2^-$. If the decision maker believes that $\rho_2^+ > \rho_2^T > \rho_1 > \rho_2^* > \rho_2^-$ and $\xi_2^+ > \xi_1 > \xi_2^- > \xi_2^*$ are satisfied, the total payoff of the low-type decision maker when he chooses T (or NT) and P is

$$B_{1,T} - \frac{C_1}{\beta_L} + \delta b_{2,T}^>$$

$$(\text{or } B_{1,NT} - \frac{C_1}{\beta_L} + \delta b_{2,T}^>).$$

Similarly, the total payoff of choosing T (or NT) and NP is

$$b_{1,T} + \delta[\lambda_s A_2^> + (1 - \lambda_s)\{\lambda_{ss} b_{2,T}^> + (1 - \lambda_{ss}) A_2^>\}]$$

$$(\text{or } b_{1,NT} + \delta[\lambda_s A_2^> + (1 - \lambda_s)\{\lambda_{ss} b_{2,T}^> + (1 - \lambda_{ss}) A_2^>\}]).$$

If δ is greater than

$$\frac{b_{1,NT} - B_{1,T} + \frac{C_1}{\beta_L}}{\{1 - \lambda_{ss}(1 - \lambda_s)\}(b_{2,T}^> - A_2^>)},$$

the low-type decision maker chooses T and P rather than NT and NP . But if the decision maker considers his first period sufficiently important, i.e. he has low enough δ , he chooses NT and NP . The probability of choosing NT and NP increases with his wrong belief (i.e., that he will forget that he chose NP and believes that he chose P). Similarly, the critical value of δ ,

$$\frac{b_{1,NT} - B_{1,T} + \frac{C_1}{\beta_L}}{\{1 - \lambda_{ss}(1 - \lambda_s)\}(b_{2,T}^< - A_2^<)},$$

⁶There are three cases in ρ - $\rho_2^+ > \rho_1 > \rho_2^- > \rho_2^T > \rho_2^*$, $\rho_2^+ > \rho_1 > \rho_2^T > \rho_2^- > \rho_2^*$, and $\rho_2^+ > \rho_1 > \rho_2^T > \rho_2^* > \rho_2^-$ under four different cases in ξ - $\xi_2^+ > \xi_1 > \xi_2^- > \xi_2^*$, $\xi_2^+ > \xi_1 > \xi_2^* > \xi_2^-$, $\xi_2^+ > \xi_2^* > \xi_1 > \xi_2^-$ and $\xi_2^* > \xi_2^+ > \xi_1 > \xi_2^-$.

is determined in the case where his updated perception of reputation is less than the critical value, i.e. $\xi_2^* > \xi_2^+ > \xi_1 > \xi_2^-$.

Proposition 6 *Under the condition that both $\rho_2^+ > \rho_2^T > \rho_1 > \rho_2^* > \rho_2^-$ and $\xi_2^+ > \xi_1 > \xi_2^* > \xi_2^-$ are satisfied, there are*

$$\delta^{***} = \frac{b_{1,NT} - b_{1,T}}{\lambda_s \lambda_f (A_2^< - A_2^>)}$$

and

$$\delta^{****} = \frac{b_{1,NT} - B_{1,T} + \frac{C_1}{\beta_L}}{\{1 - \lambda_{ss}(1 - \lambda_s)\}(b_{2,T}^> - A_2^>)}$$

such that

1. the low-type decision maker chooses NT and NP in the first period if $\delta \leq \delta^{***}$ (with $\delta^{***} \leq \delta^{****}$) or $\delta \leq \delta^{****}$ (with $\delta^{****} < \delta^{***}$),
2. the low-type decision maker chooses T and NP in the first period if $\delta > \delta^{***}$ (with $\delta^{***} \leq \delta^{****}$), or
3. the low-type decision maker chooses T and P in the first period if $\delta > \delta^{****}$ (with $\delta^{****} < \delta^{***}$) (See Appendix E).

The low-type decision maker chooses NT and NP if he considers his current period sufficiently important. In the case of $\rho_2^+ > \rho_2^T > \rho_1 > \rho_2^* > \rho_2^-$, the value function falls more if the decision maker chooses NP than if he chooses P given a fixed perception of reputation. So, the change in the perception of reputation plays an important role in this case. If δ is greater than δ^{***} (with $\delta^{***} \leq \delta^{****}$), the low-type decision maker chooses T and NP rather than NT and NP because the value function increases with the updated perception of reputation. But, if δ is less than δ^{***} (with $\delta^{***} \leq \delta^{****}$) since he considers his first period sufficiently important, he chooses NT and NP . If δ is greater than δ^{****} (with $\delta^{***} > \delta^{****}$), since the increase in the value function from updated self-reputation can compensate for the decrease in the value function from the perception of reputation when the decision maker chooses T and P rather than NT and NP , he chooses T and P in the first period. Similarly, the critical value of δ if the decision maker believes that

$\rho_2^+ > \rho_2^T > \rho_1 > \rho_2^* > \rho_2^-$ and $\xi_2^+ > \xi_2^* > \xi_1 > \xi_2^-$ is determined. By considering other cases,⁷ I conclude that the low-type decision maker whose ex-post payoff of choosing *CH* in the second period is sufficiently low never chooses *NT* and *P*, but chooses *NT* and *NP*, *T* and *NP*, or *T* and *P* by following the critical value of δ , prior beliefs and the updated beliefs.

4 Conclusion

I investigated questions related to the self-control of individuals who have imperfect information about their ability to stop bad habits and to keep a resolution. I extend the model of Benabou and Tirole (2004) by examining how a decision maker's perception of his reputation (i.e., his belief regarding what others think of him), in addition to self-reputation, affects his choices.

A two-period model is developed in this paper. Each period is divided into two sub-periods. There are three choices in the first sub-period: continuing a habit, telling friends about one's resolution only after choosing not to continue the habit, and not telling friends after choosing not to continue the habit. In the second sub-period, the decision maker chooses whether to persevere or not. In the second period, since it is the last period and thus there is no reputation to protect, the decision maker who has a strong ability to keep a resolution in the stress period (the high type) chooses to persevere while the decision maker who has a weak ability to keep a resolution (the low type) chooses not to persevere. The critical value which guarantees that the decision maker chooses to tell his friends is determined, as well as the one which guarantees that the decision maker chooses to make a resolution in the second period.

The high-type decision maker always chooses to persevere if he chooses not to continue his habit in the first period because this choice can increase his future payoff. But,

⁷Those are $\rho_2^+ > \rho_2^T > \rho_1 > \rho_2^- > \rho_2^*$, $\rho_2^T > \rho_2^+ > \rho_1 > \rho_2^* > \rho_2^-$, $\rho_2^T > \rho_2^+ > \rho_1 > \rho_2^- > \rho_2^*$, $\rho_2^+ > \rho_2^T > \rho_2^* > \rho_1 > \rho_2^-$, $\rho_2^T > \rho_2^+ > \rho_2^* > \rho_1 > \rho_2^-$, and $\rho_2^T > \rho_2^* > \rho_2^+ > \rho_1 > \rho_2^-$.

he may or may not tell his friends about his resolution in the first period because the perception of reputation can either increase or decrease his future payoff.

In the case of a low-type decision maker, the paper particularly focuses on the decision maker who makes a resolution in the first period but does not persevere in the second period. By fixing these two choices, the other four choices are identified. One of the most important observations regarding the low types is that higher self-reputation is not always better. The critical discount rate which guarantees a change in the choices (telling his friends and not to persevering, telling his friends and persevering, and not telling his friends and not persevering) is examined. Under some conditions, the low-type decision maker always chooses not to tell his friends and not to persevere after choosing not to continue his habit in the first period. However, he sometimes chooses not to tell his friends and not to persevere when he considers his first period sufficiently important. Finally, the low-type decision maker never chooses not to tell his friends and to persevere because the choice can decrease both his current payoff and his future payoff.

As an extension of the model, it is possible to consider the optimal timing of telling friends. In this case, the possibility of telling a lie (even if the decision maker fails to persevere, he tells his friends that he succeeds in persevering) may be examined for each type of the decision maker if telling a lie can increase the perception of reputation and thus change future payoff. In this paper, I assume that there is a critical value of perception of reputation which guarantees changes of payoff for the decision maker. I also assume that the choice of each type of decision maker is fixed in the second sub-period of the second period. Changes in these important assumptions would lead to different results, which are left for future research.

References

Ainslie, G., 1975. Specious reward: A behavioral theory of impulsiveness and impulse control. *Psychological Bulletin*, 463-496.

- Ainslie, G., 2001. *Breakdown of Will*. Cambridge University Press.
- American Cancer Society, 2005. *Guide to Quitting Smoking*. American Cancer Society Press.
- Becker, G., Grossman, M., Murphy, K., 1994. An empirical analysis of cigarette addiction. *American Economic Review* 84, 396-418.
- Becker, G., Murphy, K., 1988. A theory of rational addiction. *Journal of Political Economy* 96, 675-700.
- Benabou, R., Tirole, J., 2002. Self-confidence and personal motivation. *Quarterly Journal of Economics* 117, 871-915.
- Benabou, R., Tirole, J., 2004. Willpower and personal rules. *Journal of Political Economy* 112, 848-886.
- Brocas, I., Carrillo, J., 2000. The value of information when preferences are dynamically inconsistent. *European Economic Review* 44, 1104-1115.
- Fudenberg, D., Levine, D., 2006. A dual self model of impulse control. Working Paper 2112, Harvard Institute of Economic Research.
- Goldman, S., 1980. Consistent plans. *Review of Economic Studies* 47, 533-537.
- Gruber, J., Koszegi, B., 2001. Is addiction rational? Theory and evidence. *Quarterly Journal of Economics* 116, 1261-1305.
- Gul, F., Pesendorfer, W., 2001. Temptation and self-control. *Econometrica* 69, 1403-1436.
- Gul, F., Pesendorfer, W., 2004. Self-control and the theory of consumption. *Econometrica* 72, 119-158.
- Gul, F., Pesendorfer, W., 2007. Harmful addiction. *Review of Economic Studies* 74, 147-172.
- Laibson, D., 1997. Golden eggs and hyperbolic discounting. *Quarterly Journal of Economics* 112, 443-477.
- O'Donoghue, T., Rabin, M., 1999. Doing it now or later. *American Economic Review* 89, 103-124.
- Peleg, B., Yarri, M., 1973. On the existence of a consistent course of action when

tastes are changing. *Review of Economic Studies* 40, 391-401.

Pollak, R., 1968. Consistent planning. *Review of Economic Studies* 35, 201-208.

Schelling, T., 1960. *The Strategy of Conflict*. Harvard University Press.

Schelling, T., 1978. Egonomics, or the art of self-management. *American Economic Review* 68, 290-294.

Schelling, T., 2006. *Strategies of Commitment*. Harvard University Press.

Strotz, R., 1956. Myopia and inconsistency in dynamic utility maximization. *Review of Economic Studies* 23, 165-180.

Thaler, R., Shefrin, H., 1981. An economic theory of self-control. *Journal of Political Economy* 89, 392-406.

Appendix A

The expected payoff of choosing not to continue his habit (*NCH*) in the second period is

$$E_2(NCH) = \gamma[\phi(\rho_2, \xi_2)\{\rho_2(B_{2,T} - C_2) + (1 - \rho_2)b_{2,T}\} + (1 - \phi(\rho_2, \xi_2))\{\rho_2(B_{2,NT} - C_2) + (1 - \rho_2)b_{2,NT}\}].$$

If $\rho_2 \geq \rho_2^T$, since the probability of telling his friends is 1,

$$E_2(NCH) = \gamma[\rho_2(B_{2,T} - C_2) + (1 - \rho_2)b_{2,T}].$$

Under the condition that the benefit when the decision maker chooses to persevere (*P*) is sufficiently high, i.e. $B_{2,T} > b_{2,T} + C_2$, the expected payoff of *NCH* is increasing with ρ_2 . Similarly, the expected payoff of *NCH* is increasing with ρ_2 if $\rho_2 < \rho_2^T$ because of

$$E_2(NCH) = \gamma[\rho_2(B_{2,NT} - C_2) + (1 - \rho_2)b_{2,NT}]$$

and $B_{2,NT} > b_{2,NT} + C_2$.

Appendix B

I examine 20 different cases by comparing ρ_1 with ρ_2^+ , ρ_2^- , ρ_2^T and ρ_2^* . Since the methods are the same for other cases, I explain only the case in which the decision maker believes $\rho_2^+ > \rho_2^* > \rho_1 > \rho_2^T > \rho_2^-$ and $\xi_2^+ > \xi_1 > \xi_2^- > \xi_2^*$. If the decision maker chooses *T* (or *NT*) and *P*, his total payoff is

$$B_{1,T} - \frac{C_1}{\beta_H} + \delta(B_{2,T}^> - C_2) \\ (\text{or } B_{1,NT} - \frac{C_1}{\beta_H} + \delta(B_{2,T}^> - C_2)).$$

Similarly, the total payoff of choosing *T* (or *NT*) and *NP* is

$$b_{1,T} + \delta[\lambda_s A_2^> + (1 - \lambda_s)\{\lambda_{ss}(B_{2,T}^> - C_2) + (1 - \lambda_{ss})A_2^>\}] \\ (\text{or } b_{1,NT} + \delta[\lambda_s A_2^> + (1 - \lambda_s)\{\lambda_{ss}(B_{2,T}^> - C_2) + (1 - \lambda_{ss})A_2^>\}].$$

The decision maker chooses to persevere regardless of telling his friends or not. Also, the decision maker chooses T and P because it is the dominant strategy for him. In each case of ρ , if either $\xi_2^+ > \xi_1 > \xi_2^- > \xi_2^*$ or $\xi_2^* > \xi_2^+ > \xi_1 > \xi_2^-$ is satisfied, T and P is the dominant strategy for the high-type decision maker.

Appendix C

The critical values of δ which guarantee the choice T and P , or NT and P , are examined in each case. Since the method of finding it is the same as for the low-type decision maker (which is the more interesting case), I will explain some possible cases here. Let us consider the case in which the decision maker believes that $\rho_2^+ > \rho_2^* > \rho_1 > \rho_2^T$ and $\xi_2^+ > \xi_1 > \xi_2^* > \xi_2^-$ are satisfied. If the decision maker chooses T (or NT) and P , his total payoff is

$$B_{1,T} - \frac{C_1}{\beta_H} + \delta(B_{2,T}^> - C_2)$$

$$(\text{or } B_{1,NT} - \frac{C_1}{\beta_H} + \delta(B_{2,T}^> - C_2)).$$

Similarly, the total payoff of choosing T (or NT) and NP is

$$b_{1,T} + \delta[\lambda_s\{\lambda_f A_2^< + (1 - \lambda_f)A_2^>\}$$

$$+ (1 - \lambda_s)\{\lambda_{ss}(B_{2,T}^> - C_2) + (1 - \lambda_{ss})A_2^>\}]$$

$$(\text{or } b_{1,NT} + \delta[\lambda_s A_2^> + (1 - \lambda_s)\{\lambda_{ss}(B_{2,T}^> - C_2) + (1 - \lambda_{ss})A_2^>\}]).$$

I first examine under what conditions the high-type decision maker chooses T and P rather than T and NP . If

$$\{1 - \lambda_{ss}(1 - \lambda_s)\}(B_{2,T}^> - C_2 - A_2^>)$$

is greater than $\lambda_s \lambda_f (A_2^< - A_2^>)$, he chooses T and P . Also, even if

$$\{1 - \lambda_{ss}(1 - \lambda_s)\}(B_{2,T}^> - C_2 - A_2^>)$$

is less than $\lambda_s \lambda_f (A_2^< - A_2^>)$, when he considers his first period sufficiently important (i.e.,

$$\delta \leq \frac{B_{1,T} - \frac{C_1}{\beta_H} - b_{1,T}}{\lambda_s \lambda_f (A_2^< - A_2^>) - \{1 - \lambda_{ss}(1 - \lambda_s)\}(B_{2,T}^> - C_2 - A_2^>)},$$

he chooses T and P . But, the high-type decision maker always chooses NT and P rather than NT and NP . Thus, the critical value of δ and the condition to make the high-type decision maker choose NT (or T) and P is determined.

Similarly, let us consider the case in which the high-type decision maker believes that $\rho_2^+ > \rho_2^* > \rho_1 > \rho_2^T$ and $\xi_2^+ > \xi_2^* > \xi_1 > \xi_2^-$ are satisfied. If the decision maker chooses T (or NT) and P , his total payoff is

$$B_{1,T} - \frac{C_1}{\beta_H} + \delta\{\lambda_f(B_{2,T}^> - C_2) + (1 - \lambda_f)(B_{2,T}^< - C_2)\}$$

$$(\text{or } B_{1,NT} - \frac{C_1}{\beta_H} + \delta(B_{2,T}^< - C_2)).$$

Similarly, the total payoff of choosing T (or NT) and NP is

$$b_{1,T} + \delta[\lambda_s A_2^< + (1 - \lambda_s)[\lambda_{ss}\{\lambda_f(B_{2,T}^> - C_2) + (1 - \lambda_f)(B_{2,T}^< - C_2)\} + (1 - \lambda_{ss})A_2^<]]$$

$$(\text{or } b_{1,NT} + \delta[\lambda_s A_2^< + (1 - \lambda_s)\{\lambda_{ss}(B_{2,T}^< - C_2) + (1 - \lambda_{ss})A_2^<\}]).$$

The decision maker always chooses T and P rather than T and NP , and always chooses NT and P rather than NT and NP . Thus, the decision maker chooses to tell his friends and to persevere if he considers his first period sufficiently important (i.e. if

$$\delta \leq \frac{B_{1,T} - B_{1,NT}}{\lambda_f(B_{2,T}^< - B_{2,T}^>)}.$$

But, if the discount rate δ is greater than the critical value, he chooses not to tell his friends and to persevere in the first period.

Appendix D

Let us first compare the total payoff of choosing T and P with that of choosing T and NP . The low-type decision maker chooses T and NP if $\delta\lambda_s(b_{2,T}^> - A_2^>)$ is less than $b_{1,T} - B_{1,T} + \frac{C_1}{\beta_L}$ and chooses T and P otherwise. Similarly, the low-type decision maker chooses NT and NP rather than NT and P if $\delta\lambda_s(b_{2,T}^> - A_2^>)$ is less than $b_{1,NT} - B_{1,NT} +$

$\frac{C_1}{\beta_L}$ and chooses NT and P if $\delta\lambda_s(b_{2,T}^> - A_2^>)$ is greater than $b_{1,NT} - B_{1,NT} + \frac{C_1}{\beta_L}$. Thus, if

$$\delta \leq \frac{b_{1,T} - B_{1,T} + \frac{C_1}{\beta_L}}{\lambda_s(b_{2,T}^> - A_2^>)},$$

I compare the total payoff between the choice T and NP and choice NT and NP . But if δ lies between

$$\frac{b_{1,T} - B_{1,T} + \frac{C_1}{\beta_L}}{\lambda_s(b_{2,T}^> - A_2^>)}$$

and

$$\frac{b_{1,NT} - B_{1,NT} + \frac{C_1}{\beta_L}}{\lambda_s(b_{2,T}^> - A_2^>)},$$

the low-type decision maker chooses either NT and NP or T and P . Finally, if δ is greater than

$$\frac{b_{1,NT} - B_{1,NT} + \frac{C_1}{\beta_L}}{\lambda_s(b_{2,T}^> - A_2^>)},$$

I compare the total payoff between choice T and P and choice NT and P . In the first area, i.e. if

$$\delta \leq \frac{b_{1,T} - B_{1,T} + \frac{C_1}{\beta_L}}{\lambda_s(b_{2,T}^> - A_2^>)},$$

the low-type decision maker always chooses NT and NP . But if

$$\frac{b_{1,T} - B_{1,T} + \frac{C_1}{\beta_L}}{\lambda_s(b_{2,T}^> - A_2^>)} \leq \delta \leq \frac{b_{1,NT} - B_{1,NT} + \frac{C_1}{\beta_L}}{\lambda_s(b_{2,T}^> - A_2^>)},$$

the low-type decision maker chooses NT and NP whenever

$$\delta \leq \frac{b_{1,NT} - B_{1,T} + \frac{C_1}{\beta_L}}{\lambda_s(b_{2,T}^> - A_2^>)} < \frac{b_{1,NT} - B_{1,NT} + \frac{C_1}{\beta_L}}{\lambda_s(b_{2,T}^> - A_2^>)}.$$

Also, he chooses T and P in the first period if

$$\delta > \frac{b_{1,NT} - B_{1,T} + \frac{C_1}{\beta_L}}{\lambda_s(b_{2,T}^> - A_2^>)}$$

in the second area. Finally, if δ is greater than

$$\frac{b_{1,NT} - B_{1,NT} + \frac{C_1}{\beta_L}}{\lambda_s(b_{2,T}^> - A_2^>)},$$

the low-type decision maker always chooses T and P . Thus, there exists the critical discount rate

$$\frac{b_{1,NT} - B_{1,T} + \frac{C_1}{\beta_L}}{\lambda_s(b_{2,T}^> - A_2^>)}$$

which guarantees the change of choices from NT and NP to T and P in the first period.

Appendix E

Under the condition that both $\xi_2^+ > \xi_1 > \xi_2^* > \xi_2^-$ and $\rho_2^+ > \rho_2^T > \rho_1 > \rho_2^* > \rho_2^-$ are satisfied, if the decision maker chooses T (or NT) and P , his total payoff is

$$\begin{aligned} & B_{1,T} - \frac{C_1}{\beta_H} + \delta b_{2,T}^> \\ & \text{(or } B_{1,NT} - \frac{C_1}{\beta_H} + \delta b_{2,T}^> \text{)}. \end{aligned}$$

The total payoff of choosing T (or NT) and NP is

$$\begin{aligned} & b_{1,T} + \delta[\lambda_s\{\lambda_f A_2^< + (1 - \lambda_f)A_2^>\} + (1 - \lambda_s)\{\lambda_{ss}b_{2,T}^> + (1 - \lambda_{ss})A_2^>\}] \\ & \text{(or } b_{1,NT} + \delta[\lambda_s A_2^> + (1 - \lambda_s)\{\lambda_{ss}b_{2,T}^> + (1 - \lambda_{ss})A_2^>\}]). \end{aligned}$$

First, I compare the total payoff of T and P with that of T and NP . If $\lambda_s \lambda_f (A_2^< - A_2^>)$ is greater than

$$\{1 - \lambda_{ss}(1 - \lambda_s)\}(b_{2,T}^> - A_2^>),$$

the low-type decision maker always chooses T and NP . But if $\lambda_s \lambda_f (A_2^< - A_2^>)$ is less than

$$\{1 - \lambda_{ss}(1 - \lambda_s)\}(b_{2,T}^> - A_2^>),$$

the decision maker chooses T and NP only when he considers his first period sufficiently important, i.e. whenever

$$\delta \leq \frac{b_{1,T} - B_{1,T} + \frac{C_1}{\beta_L}}{\{1 - \lambda_{ss}(1 - \lambda_s)\}(b_{2,T}^> - A_2^>) - \lambda_s \lambda_f (A_2^< - A_2^>)}$$

Similarly, the low-type decision maker chooses NT and NP if

$$\delta \leq \frac{b_{1,NT} - B_{1,NT} + \frac{C_1}{\beta_L}}{\{1 - \lambda_{ss}(1 - \lambda_s)\}(b_{2,T}^> - A_2^>)}$$

is satisfied, and chooses NT and P if δ is sufficiently high. Under the condition that $\lambda_s \lambda_f (A_2^< - A_2^>)$ is greater than

$$\{1 - \lambda_{ss}(1 - \lambda_s)\}(b_{2,T}^> - A_2^>) \text{ (or } \delta^{***} \leq \delta^{****}\text{),}$$

there is the critical value

$$\delta^{***} = \frac{b_{1,NT} - b_{1,T}}{\lambda_s \lambda_f (A_2^< - A_2^>)}$$

which guarantees the change of the choice from NT and NP to T and NP in the first period. Similarly, if $\lambda_s \lambda_f (A_2^< - A_2^>)$ is less than

$$\{1 - \lambda_{ss}(1 - \lambda_s)\}(b_{2,T}^> - A_2^>) \text{ (or } \delta^{****} > \delta^{***}\text{),}$$

there is the critical value

$$\delta^{****} = \frac{b_{1,NT} - B_{1,T} + \frac{C_1}{\beta_L}}{\{1 - \lambda_{ss}(1 - \lambda_s)\}(b_{2,T}^> - A_2^>)}$$

such that if $\delta \leq \delta^{****}$, the low-type decision maker chooses NT and NP , and if $\delta > \delta^{****}$, the decision maker chooses T and P in the first period.

Individual researchers, as well as the on-line and printed versions of the CERGE-EI Working Papers (including their dissemination) were supported from the following institutional grants:

- Economic Aspects of EU and EMU Entry [Ekonomické aspekty vstupu do Evropské unie a Evropské měnové unie], No. AVOZ70850503, (2005-2010);
- Economic Impact of European Integration on the Czech Republic [Ekonomické dopady evropské integrace na ČR], No. MSM0021620846, (2005-2011);

Specific research support and/or other grants the researchers/publications benefited from are acknowledged at the beginning of the Paper.

(c) Jung Hun Cho, 2007.

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical or photocopying, recording, or otherwise without the prior permission of the publisher.

Published by
Charles University in Prague, Center for Economic Research and Graduate Education (CERGE)
and
Economics Institute ASCR, v. v. i. (EI)
CERGE-EI, Politických vězňů 7, 111 21 Prague 1, tel.: +420 224 005 153, Czech Republic.
Printed by CERGE-EI, Prague
Subscription: CERGE-EI homepage: <http://www.cerge-ei.cz>

Editors: Directors of CERGE and EI
Managing editors: Deputy Directors for Research of CERGE and EI

ISSN 1211-3298
ISBN 978-80-7343-142-6 (Univerzita Karlova. Centrum pro ekonomický výzkum
a doktorské studium)
ISBN 978-80-7344-131-9 (Národohospodářský ústav AV ČR, v. v. i.)



CERGE-EI
P.O.BOX 882
Politických vězňů 7
111 21 Praha 1
Czech Republic
<http://www.cerge-ei.cz>