A Principal-Agent Model with Ego in the Utility Function

Eugenio Giolito

August 2001

Abstract

I develop a simple principal agent model witah an Ego "component" in agentt's utility function. The main result is that, when ability and effort are perfect substitutes in a job and the principal has superior information, it will be optimal for the principal to make the agent believe that he is high ability no matter whether if is true or not. That result is compatible with many firms "motivation" policies

1. Introduction

The idea I have in mind is to look for an alternative explanation to several facts we can see everyday. For example, we can observe in the Department that some students are doing a big effort looking for a good idea for their thesis, maybe pursuing future academic jobs. Other people are more "practical": after finishing second year they look for the fastest way to get out of here and get a well paid job. Assuming that academic jobs are not necessarily better paid than government or private sector jobs, applying a standard model of human capital accumulation will not help much to understand this issue. One explanation could be that they like what they do, but it seems that there must be something else. Getting a job as an Assistant Professor in a well-recognized University is a signal of academic success that is not easy to achieve for a lot of Ph.D. Students, and maybe this could be an explanation. Some of these students, when they took the decision to pursue a Ph.D., left a job, a house, a town, even a country to undertake a project with a very uncertain return. They undertake a project that could finish if they

fail an exam for the second time after the first year. Are they risk lovers? I think we can go farther.

Another story that could be interesting to think about is, for example, the behavior of some Wall Street executives, or owners of firms that make millions and millions and never retire. Can we explain Bill Gates behavior with a standard profit-maximizing framework? Is Microsoft really trying to maximize profits in this complex environment? Or maybe is trying to achieve monopoly whatever it costs? Are the Wall Street guys just looking for more money? Maybe we can view high salaries, high profit, or high market share as a signal of success in a highly competitive business environment. Are they trying to earn or trying to win?

A third story is the most simple. In any office I ever worked there was always the "workaholic" guy, who always comes to work early and is the last that goes home. Is the typical employee one who comes to work even when he is sick, and says, "How would the job get done if I stay home?" This "workaholic" employee can be the boss, a secretary or even the guys who serves coffee. However important or not his job is, this guy feels that the firm cannot survive if he is not there, taking decisions or serving coffee, because he is the best in what he is doing. For this kind of worker, a little act of public recognition from the boss, or even his peers (like, you are the best handyman I've ever seen), could be more important than a salary raise (unless the salary raise is only for him, of course).

What have these stories in common? Good grades, faculty position and tenure in the academic career, money or market share in the business world and public recognition in the case of the simple employee are signals of success, signals that people are good in what they do. Not only signals for others, but also signals for themselves. Getting these signals, people discover how good they are, and people like to discover they are good. What I would like to do is to take into account this "like to be good" or "ego" characteristic as a component in the utility function and try to analyze its effects in job related relationships in a principal-agent framework. In other words, I will try to internalize the competitive nature of human beings with an ad-hoc utility function and discuss the effects of this characteristic in everyday professional interrelationships.

The main result of this paper is that, when ability and effort are perfect substitutes in a job and the principal has superior information about the nature of the job and agent's ability, it will be optimal for the principal to reinforce the agent's Intrinsic Motivation (that is, the idea that people might be willing to exert effort for reasons other than salary). That is, make the agent believe that he is high ability no matter whether if is true or not. This result is very compatible

2

with many firms policies about motivation of the employees..

The plan of the paper is the following:

Section 2 is an overview of the very recent literature in Ego Utility and Instrinsic Motivation in Economics. I will try also to explain the particular nature of the Utility function I will use.

In Section 3 I will focus on the simplest example of the workaholic employee and analyze the potential incentives of "untruthful" Intrinsic Motivation that the employers can have when they face these kind of agents. After developing a very simple model, I will discuss the possibility that the principal could have incentives to send positive (and costless) signals about the agent's ability. In this way the principal will get more effort from the agent without increasing her salary (trough Intrinsic Motivations). These signals could be simple congratulations, fancy position names (like Vice President of Incoming Calls, or something like that) or costless privileges like "executives bathroom",etc.

In Section 4 I try to point out steps for my future research.

2. The Issue

2.1. Overview of the Recent Literature

The literature about Ego Utility and Intrinsic Motivation is very recent in Economics. There are three papers very related with the objective of this work: Köszegi (2000a and 2000b) and Bernabou and Tirole (2000).

Köszegi (2000a and 2000b) develops an explicit model accounting for Ego Utility. He assumes that the agents derive utility from positive views about the self, holding constant standard utilitarian outcomes usually relevant in Economics. The model is quite general, in order to fit several of economic applications. Köszegi's agents voluntary collect information about their ability, and they can decide when to stop this learning process. One of the key finding in his paper is that individuals tend to acquire overconfidence beliefs on average. The driving force in his model is, as Köszegi says:

> "the agent's desire to affect marginal changes in her beliefs, that is, to manipulate the choice of tasks to manage the signals she receives about herself. There are two reasons why the agent may choose to do this. Firstly, when she is satisfied with her present beliefs, she might distort her instrumental choices to avoid receiving information about

3

herself. I call this the self-image protection motive. On the other hand, if she is dissatisfied with her current perception of herself, she might go out of her way to try to improve her beliefs. This is the self-image enhancement motive."

There are two important differences between the models above and the study I would like to pursue here. First, I will jump directly into applications, analyzing a principal agent model. Second, my "agent" is less sophisticated. She cannot manage the learning process, and receives signal passively, updating her beliefs about her ability once a exogenous signal is received. In line with Bernabou and Tirole (2000), I study the interaction of individuals that are fully rational and assuming, as they say

" the individual is an information processor who extracts from his environment signals that are relevant for his self-confidence. Although people surely make mistakes in processing information, we account for the fat that they cannot systematically and repeatedly fool themselves, or others for that matter"

Even though Bernabou and Tirole also analyze a principal-agent model where the principal has superior information, my approach will differ form theirs in two aspects. On the one hand, I will try to study an environment where self-confidence about ability is relevant even if ability and effort are perfect substitutes in the production function (they consider nonseparability essential in their argument).. On the other hand, I will concentrate on "intrinsic motivation" rather than interaction between extrinsic (salary) and intrinsic motivation, as they do in their paper.

2.2. The Utility Function

Our agent has a utility function with the same spirit as in Köszegi (2000a and 2000b), and takes the form:

$$V(w, e, \theta) = \begin{array}{ll} U(w, e) + \Phi\left(Q(\theta, e)\right) & if \quad \theta \geq \theta^*, \ \theta^* \ given \\ U(w, e) & otherwise \end{array}$$

where Q is the production function of the agent, w is salary, θ is ability and e effort.

$$U_w > 0 \qquad U_e < 0 \qquad \Phi_e > 0$$
$$U_{ww} \leq 0 \qquad U_{ee} < 0$$

4

$\|U_e\| > \Phi_e$

As one can see, this function is partially a "traditional" Utility function concave in consumption and with increasing disutility in effort, on the one hand, and convex in effort, on the other hand, if ability reaches a certain level. The intuition of this function is that, if the guy thinks his ability is higher than a certain level (he think he is good in what he does), doing his job provides an additional amount of utility independent of the salary or profit he receives (that we can call "ego rents"). The intuition of including the value of output as a variable in the utility function is quite simple. What I am trying to analyze here is not the effect of "enjoying the job" in the utility function. I would rather internalize the effect of "doing an important job" in the utility function and the value of the output could be a reasonable approximation of that idea.

Differentiating the utility function, for a given level of ability, we get

$dV(w, e, \theta) = U_w dw + U_e de + \Phi_e de = 0$

$MRS = \frac{dw}{de} = -\frac{[\Phi_e + U_e]}{U_w} > 0$

The slope of the MRS with the presence of "ego rents" is lower than in the normal case, so the agent is willing to accept a higher increase in effort for the same increase in consumption than in a traditional utility function. That assumption will be crucial in the analysis of the agent behavior. I assume that the agent does not know his true level of ability. He makes beliefs over signals he receives and updates his beliefs using Bayes rule. How these signals are generated will depend on the case we want to analyze. For the purpose of this paper, I will assume that the signals are chosen by the principal in order to maximize profit.

3. The model

3.1. Assumptions

- The workaholic guy (from now on, the "agent") has an ability level of $\theta \in [0, 1]$ with distribution function $F(\theta)$

- The production Function of the agent takes the form

(1) $Q(\theta, e) = \alpha\theta + \beta e$

where θ is ability and e is effort

$0 \leq \alpha \leq 1$

$0 < \beta < 1$

These assumptions mean that the Output can depend only on effort, but never only on ability. As we can see, ability and effort are perfect substitutes. I think that using this production form will help us better to understand the incentives of the principal to manipulate the agent's beliefs about his own ability.

- The Utility function of the Agent takes the form

$$(2) \quad V(w, e, \theta) = \begin{array}{ll} w - \frac{e^2}{2} + \frac{1}{2}\left(Q(\theta, e)\right)^2 & if \quad \theta \geq \theta^*, \; \theta^* \; given \\ w - \frac{e^2}{2} & otherwise \end{array}$$

- $V(0, e, \theta) = 0 \forall e, \theta$

  I make here the assumption that if, the consumption is 0, the level of Utility is 0 (people cannot live on Ego)

- The principal maximizes net profits

$(3) \quad \pi(\theta, e) = \alpha\theta + \beta e - w$

- Profits are observable and verifiable.

- At this stage I will omit the moral hazard problem from the agent and assume that effort is observable by the principal.

- The principal has full information. She knows the true value of θ, and the parameters of the Production function,α and β.

- The agent does not know the true values α and β

- The agent only knows the Distribution Function $F(\theta)$, from which he makes his prior beliefs about his level of ability θ.

- The minimum level of ability required by the agent to enjoy the "ego utility", θ^*, is exogenous to the model.

- The prior beliefs about the level of ability θ are updated by the agent, using Bayes Rule. The agent receives a signal A, sent by the principal, about the value of the parameter in the utility function concerning his ability.

- In a maybe excessive simplification, I will assume that the agent takes A and b as given (as if they were the true parameters), updating beliefs only over θ given the signal A received from the principal. After receiving the signal A, the agent updates his believes according to the conditional distribution $G(\theta/_A)$,and conditional density $g(\theta/_A)$

- For $A_1$and A_2, with $A_1 > A_2$, $G(\theta/_{A_1}) \leq G(\theta/_{A_2})$for all $\theta \in [0,1]$. That is the distribution of ability for a higher signal of the ability parameter A_1 in the production function FOSD's the distribution of ability for a lower signal A_2. The intuition of this assumption is as follows: the greater is the level of the parameter concerning the ability in the production functions, the greater the probability that the agent hired to perform that task is high ability, given that the level of ability is observed by the principal.

The purpose of the rest of the section is to analyze the incentives of the principal to announce the true value of the parameter, that is, to announce $A = \alpha$. In other words, I would like to show that, in the context that finding out the ability of the agent is costless for the principal, it will be dominant for her to tell the agent that he has more ability than he really does. We are assuming that also sending signals is costless (anything else than salary).

Why I assume that α and β are private information of the principal, is something that is worthy to point out. Even though this is a one period model, I am trying to represent a long lasting relationship. If the agent knows the true value of α and β, he will know, when profits are realized, the true value of θ. If the principal has any incentives to conceal this information, she will be able to do so only for one period. I think in this way we can think about a multiperiod model without need to explicitly model it.

3.2. The Interaction Between Principal and Agent

1. The owner (principal) offers the employee a wage compensation w

2. The principal announces the parameters A and b, chosen optimally and that not necessarily coincide with the true values α and β

3. The agent accepts or rejects this schedule. If the agent rejects, the interaction is over. The employee earns his reservation value $\bar{U}$ (assume for simplicity $\bar{U} = 0$)

7

4. The agent chooses the level of effort $e \in [1, \infty)$, taking A and b as given

5. Profits are realized

Expected payoffs of both agent and principal, if employee accepts a wage w, depends on the level of effort e and ability θ. The agent's expected payoffs are:

$$w - \frac{e^2}{2} + \left[Prob(\theta \geq \theta^*/_A) \frac{1}{2} \left[A \left[E(\theta/_{\theta \geq \theta^*, A}) \right] + be \right]^2 \right]$$

$$= w - \frac{e^2}{2} + \left[(1 - G(\theta^*/_A)) \left(\int\limits_{\theta^*}^{1} \frac{1}{2} [A\theta + be]^2 \, g(\theta/_A) d\theta \right) \right]$$

If the agent accepts the wage w, she will choose a level of effort $\widehat{e}\,(w)$ where

$$\widehat{e}\,(w) = \underset{e}{argmax} \left[w - \frac{e^2}{2} + (1 - G(\theta^*/_A)) \left(\int\limits_{\theta^*}^{1} \frac{1}{2} [A\theta + be]^2 \, g(\theta/_A) d\theta \right) \right]$$

Consequently, she will accept the contract if

$$\left[w + (1 - G(\theta/_A)) \left(\int\limits_{\theta^*}^{1} \frac{1}{2} [A\theta + be]^2 \, g(\theta/_A) d\theta \right) \right] \geq \frac{e^2}{2}$$

Therefore, the principal faces the following optimization problem:

$$\underset{w,e,A,b}{max} \, [\alpha\theta + \beta e - w]$$

$$s.t.$$

(4) $\left[w - \frac{e^2}{2} + (1 - G(\theta/_A)) \frac{1}{2} \left[A \left[E(\theta/_{\theta \geq \theta^*, A}) \right] + be \right]^2 \right] \geq 0$ (Individual Rationality (IR))

and

(5) $\alpha\theta + \beta e = A\theta + be$

The second restriction makes sense if we consider that profits are observable. The principal can announce parameters that differ of the true values but with the restriction that the linear equation with the announced values should match the profit level.

3.3. Full Information Case

Assume that the true values of θ, α and β are known by the agent. In this case, the principal's problem is:

$$\max_{w,e} \pi(\theta, e) = [\alpha\theta + \beta e - w]$$

$s.t.$
$$w - \frac{e^2}{2} + \frac{1}{2}[\alpha\theta + \beta e]^2 \geq 0 \qquad \text{(Individual Rationality (IR))}$$

the Lagrangian is
$$L = \max_{w,e} [\alpha\theta + \beta e - w] + \lambda \left(w - \frac{e^2}{2} + \frac{1}{2}[\alpha\theta + \beta e]^2 \right)$$

- FOC_w

$$-1 + \lambda = 0 \implies \lambda = 1$$

- FOC_e

$$\beta + \lambda \left[-e + \alpha\theta\beta + \beta^2 e \right] = 0$$

arranging terms we get

$$\text{(6)} \quad e^* = \frac{\beta(1 + \alpha\theta)}{\left(1 - \beta^2\right)}$$

Where we can see that the optimal effort exerted by the agent will depend positively on his level of ability. Notice that here, even the separability in the production function, we get some kind of "complementarity" , from the "ego utility" of the agent

- By FOC_λ :

$w^* = \frac{e^2}{2} - \frac{1}{2}[\alpha\theta + \beta e]^2$

The salary will be such that the agent earns his reservation value of 0. Notice that the wage less than compensate the cost of effort. The second term in the RHS should be interpreted as the "ego rents" that the principal extract from the agent. Combining the FOCs we get that the optimal wage will be

(7) $w^* = \frac{\beta^2 - \alpha^2\theta^2}{2(1-\beta^2)}$

This apparently counter-intuitive result tell us that the equilibrium wage in the full information problem will be decreasing in the agent's ability and in the ability parameter of the Production function. The intuition of this result comes from the linearity of the production function and from the "ego utility". Ability and effort are substitutes, so the high ability individual has a "sunk cost". He is not able to choose his level of ability and, even worse, he enjoys that ability.

3.4. Hidden Information Case

As I said above, here the principal's problem is:

$$\max_{w,e,A,b} \pi(\theta,e) = [\alpha\theta + \beta e - w]$$

$s.t.$ $(4), (5)$

The Lagrangian is

$$L = \max_{w,e,A,b} [\alpha\theta + \beta e - w] + \lambda\left[w - \frac{e^2}{2} + (1 - G(\theta^*/_A))\frac{1}{2}\left[A\left[E(\theta/_{\theta \geq \theta^*},_A)\right] + be\right]^2\right]$$
$$+\gamma\left[\alpha\theta + \beta e - A\theta + be\right]$$

- FOC$_w$

$-1 + \lambda = 0 \Longrightarrow \lambda = 1$

- FOC$_A$

$(1 - G(\theta^*/_A))\left([AE(\theta/_{\theta \geq \theta^*},_A) + be]\left[E(\theta/_{\theta \geq \theta^*},_A) + A\frac{d}{dA}E(\theta/_{\theta \geq \theta^*},_A)\right]\right)$
$-G_A(\theta^*/_A)\frac{1}{2}[AE(\theta/_{\theta \geq \theta^*},_A) + be]^2 = \theta\gamma$

- FOC$_b$

10

$(1 - G(\theta^*/_A))[A(E(\theta/_{\theta \geq \theta^*, A})) + be] = \gamma$

Combining the last two FOCs we get

(8) $\left\{ E(\theta/_{\theta \geq \theta^*, A}) \left[1 - \frac{1}{2}\frac{G_A(\theta^*/_A)}{(1-G(\theta^*/_A))}A\right] - \theta + A\left(\frac{d}{dA}E(\theta/_{\theta \geq \theta^*, A})\right) \right\} = \left(\frac{G_A(\theta^*/_A)}{(1-G(\theta^*/_A))}\frac{1}{2}\right) be$

- FOC_e

  $\beta + (1 - G(\theta^*/_A))[Ab(E(\theta/_{\theta \geq \theta^*, A}) + b^2 e] + \gamma[\beta - b] = e$

Substituting γ from FOC_b into FOC_e and rearranging terms

- $\beta + (1 - G(\theta^*/_A))[A(E(\theta/_{\theta \geq \theta^*, A})\beta + \beta be] = e$

(9) $e^*(\beta, A, b) = \frac{\beta\left(1 + \left[(1-G(\theta^*/_A))A\left(E(\theta/_{\theta \geq \theta^*, A})\right)\right]\right)}{[1-(1-G(\theta^*/_A))b\beta]}$

As the equation shows, the optimal level of effort   from the point of view of the principal is increasing in true parameter β (the slope of the production function with respect to effort), and fundamentally in A, the announced value of the ability parameter. Obviously, it also depends on b,but this relationship is not crucial at all. Notice that effort has no relationship with the true parameter of ability θ.

The dependence of effort on A comes from three different sources:

a) Directly via the "ego utility" that depends on the reported parameter A of the production function

b) The beliefs about the level of ability, increasing in the reported parameter A. As the reported ability parameter increases, that makes the numerator to increase and the denominator to decrease.

c) The conditional expectation of ability, also increases in the announced value of the parameter corresponding to ability in the production function.

FOC_γ
$[\alpha\theta + \beta e - A\theta - be] = 0$

11

- By FOC$_\lambda$

$w = \frac{e^2}{2}\left[1 - (1 - G(\theta/A))\, b^2\right] - (1 - G(\theta/A))\frac{1}{2}\left\{A^2\left[E(\theta/_{\theta \geq \theta^*},A)\right]^2 + A\left[E(\theta/_{\theta \geq \theta^*},A)\right]be\right\}$

Substituting e^* from (9) into FOC$_\lambda$

(10) $w^* = \frac{1}{2}\beta^2\dfrac{\left[1+\left((1-G(\theta^*/A))AE(\theta/_{\theta \geq \theta^*},A)\right)\right]^2}{\left[1-(1-G(\theta^*/A))b\beta\right]}\dfrac{\left[1-(1-G(\theta/A))b^2\right]}{\left[1-(1-G(\theta^*/A))b\beta\right]}$

$\qquad - (1 - G(\theta/A))\frac{1}{2}AE(\theta/_{\theta \geq \theta^*},A)\left[AE(\theta/_{\theta \geq \theta^*},A) + be\right]$

Define $\Psi(b) = \dfrac{\left[1-(1-G(\theta/A))b^2\right]}{\left[1-(1-G(\theta^*/A))b\beta\right]}$

Notice that $\Psi < 1$ when $b < \beta$, $\quad \Psi'(b) < 0$

So we can express (10)

$w^* = \frac{1}{2}\beta^2\dfrac{\left[1+\left((1-G(\theta^*/A))AE(\theta/_{\theta \geq \theta^*},A)\right)\right]^2}{\left[1-(1-G(\theta^*/A))b\beta\right]}\Psi(b) - (1 - G(\theta/A))\frac{1}{2}AE(\theta/_{\theta \geq \theta^*},A)\left[AE(\theta/_{\theta \geq \theta^*},A) + be\right]$

Using (3), FOC$_\gamma$, (9) and (10) we can reach an expression for profits as a function of parameters A and b, given α, β and θ.

(11) $\pi(A,b) = \left[\alpha\theta + \beta e^*(A,b)\right]\left\{1 + \frac{1}{2}\left[(1 - G(\theta/A))\,AE(\theta/_{\theta \geq \theta^*},A)\right]\right\}$

$\qquad\qquad -\frac{1}{2}\beta^2\left[e^*(A,b)\right]^2\Psi(b)\left[1 - (1 - G(\theta^*/A))\,b\beta\right]$

It is easy to see that profits are increasing in the announced parameter of ability A, and slightly decreasing in the announced parameter of effort b. Therefore, the principal will announce an ability parameter higher than the true value to the agent. Telling the agent that the job he is performing is for highly qualified people, the principal makes the agent exert more effort. And here our Workaholic employee, feeling smarter than he really is, let the principal extract the "ego rents" and increase her benefits.

4. Conclusion and Ideas for Future Research

Under very specific assumptions, I found that manipulating the agent's beliefs in order to increase her "intrinsic motivation' could be an optimal policy for firms. My purpose for the future is to try to generalize this result for the case where

effort is not observable, and what is most interesting, the case where ability really matters in the production function.

Take the case of the Ph.D. Student. Here the Agent (the University) maximizes the output (thesis), using scarce resources (financial support, faculty's time, etc.) to help students in their dissertations. In order to distribute these scarce resources efficiently the University designs a screening mechanism (a device that could be an exam) to determine which student should get support. This mechanism sends a signal to both principal and agent. Here both are interested to know the true value of θ. I am thinking of a common value model, where, for example, the student makes updates about her prior beliefs (originated in the device) from the receptivity she receives from professors. If her beliefs about being good tend to zero (so the Utility function is the usual concave function), so she will try to get the degree in the easiest and fastest way possible. If the updating mechanism makes her believe she has a great potential, she will increase the effort to do an outstanding dissertation. What would be also an interesting point here is the mechanism design problem from the point of view of the principal. The principal is aware of the student's utility function and, if a "wrong" signal make her to believe she is not good enough, the principal will not be able to obtain her best output (e.g. an outstanding thesis even if she is able to do so)

References

- Bernabou, R. and Jean Tirole (2000):"Self-Confidence and Social Interactions", NBER Working Paper 7585

- Köszegi, Botond (2000a):"Ego Utility, Overconfidence, and Task Choice", mimeo, UC Berkeley

- Köszegi, Botond (2000b):"Ego Utility and Information Acquisition", mimeo, UC Berkeley

- Mas-Colell, Whinston and Green: Microeconomic Theory. Oxford University Press 1995