Dynamic programming approach to the robust principal-agent problem

Sheng, Shunan

2020

Sheng, S. (2020). Dynamic programming approach to the robust principal-agent problem. Student Research Paper, Nanyang Technological University, Singapore. https://hdl.handle.net/10356/151274

https://hdl.handle.net/10356/151274

Dynamic Programming approach to The Robust Principal-agent Problem

Sheng Shunan
School of Physics and Mathematical Sciences

Asst. Prof. Patrick Pun Chi Seng
School of Physics and Mathematical Sciences

***Abstract*— Principal-agent models are studied to incorporate the moral hazard where the agent has unobservable behavior. This paper considers a special formulation of the principal-agent problem with finite time lump-sum payment, which can be interpreted as the two-player stochastic differential game. Inspired by the latest works, we exploited the dynamic programming approach to solve the stochastic control problem. In addition, we investigated the robust formulation by introducing uncertainty to the drift function, resulting in an inaccurate observation about the output process.**

Key Words: Principal-agent problem, HJBI equation, robust control

I. INTRODUCTION

Principal-agent problem studies the optimal contract between two parties—a principal and an agent, the problem arises when the agent and principal have conflicting interests and the agent has more information than the principal, i.e. the agent's own efforts. In this case, the principal cannot guarantee that the agent is performing to realize the principal's best interest [Bebchuk and Fried(2004)] . For example, when the agent's actions are unobservable and the agent's efforts that are achieving the goal of the principal are costly to the agent, finding an optimal contract will be of great importance to study.

As a classical moral hazard problem in microeconomics, the principal-agent problem has wide applications in various fields of economics and finance, such as, in insurance and portfolio management [Patrick Bolton(2005)]

Systematic efforts have been invested in solving the moral hazard since the 1960s. In 1970, Zeckhauser [Zeckhauser(1970)] proposed the first formal model of moral hazard for the insurance policy, with solutions being computed using the first-order conditions. This method, however, was criticized by Mirrlees[Mirrlees(2017)] to be restrictive and even generally inapplicable. In 1987, Holmström and Milgrom [Holmstrom and Milgrom(1987)] pioneered the study on principal-agent problems in continuous time with CARA utility functions and proved that if the agent's effort influences the drift of the output process, but not the volatility, then the optimal contract is linear. In the 2000s, researchers started to realize the dynamic nature of the principal-agent problem and proposed multiple approaches exploiting the dynamic programming principle [Sannikov(2008)]. However, not until 2018, a systematic dynamic programming approach had been finally developed by Cvitanić in his remarkable work [Cvitanić(2018)], on which this paper further generalizes the approach to the robust formulation of the problem.

Indeed, understanding the dynamic nature is central to solve the principal-agent problem. Reminiscent of the Cvitanić et al., 2018, this paper applies the dynamic programming approach by optimizing the reduced contract for the principal's objective value function. Our main contribution is the following: we generalize the dynamic programming approach to the robust formulation of the principal-agent problem with the volatility control. The robustness comes from the uncertainty in the drift function, influencing the output process by adding an extra variable. As a result, the HJB equation is replaced by the HJBI equation indicating that we maximize the worst-case situation given the uncertainty.

II. LITERATURE REVIEW

A. Dynamic programming

Exploiting the dynamic nature of the principal-agent problems receive prevalence among researchers. Sannikov [Sannikov(2008)] studies the infinitesimal decomposition structure of the agent's dynamic value function induced by the dynamic programming principle. Inspired by Sannikov's work, Civatanci [Cvitanić(2018)] restrict the family of admissible controls to a tailored subset for which the agent's value process allows a dynamic programming presentation. Remarkable, they conclude that, under mild conditions, the supremum of the principal's problem over the subfamily is equal to the supremum over all feasible contracts. Undoubtedly, Sannikov's work offers a systematic and almost perfect answer to the principal-agent problem with the lump payment. However, in most cases, using the dynamic programming approach leaves the problem analytically intractable so that researchers must resort to the Monte Carlo simulation at the end.

B. Robust optimization

The concept of robust optimization has come with several approaches to protecting the decision-maker against parameter ambiguity and stochastic uncertainty[Gabrel et al.(2014)Gabrel, Murat, and Thiele]. Among its wide applications, robust portfolio optimization is out of paramount importance to business practitioners. Guo [Guo et al.(2019)Guo, Langrené, Loeper, and Ning] studies the problem of utility maximization with an uncertain covariance matrix by translating it into a two-player stochastic differential game. Similarly, Maenhout [Maenhout(2004)] proposes a robust approach to the dynamic portfolio and consumption problem of an investor who concerns the model uncertainty and formulates the problem with max-min expected utility. For the principal-agent problem, on the other hand, incorporating the uncertainty in the drift process is a reasonable attempt to cope with the dynamic environment as well, under which both the agent and the principal have uncertainty preferences.

III. Principal-agent Problem

A. The benchmark setting

Given the control process $\nu = (\alpha, \beta)$, a $\mathrm{F}-optional$ process with values in $A \times B$ for some subsets A,B of finite dimensional spaces. The reference-controlled state equation is defined by the stochastic differential equation, driven by an n-dimensional P-Brownian motion $\mathrm{W}_t^{\mathrm{P}}$, with drift and variance function controlled by α, β, respectively,

$$dX_t = \sigma(X, \beta_t)\left(\lambda_t(X, \alpha_t)\,dt + dW_t^P\right) \tag{1}$$

Given a contact payment $C_T = F(X_T)$ that is a function of the terminal value X_T, the agent's and principal's utility are denoted as $U_A(C_T), U_P(X_T - C_T)$. Next, we introduce the agent's effort function $c_t(\nu_s)$, which is unobservable to principal and the constant reservation utility R. The agent's problem turns out to be,

$$V_A(X_T) = \sup_{\nu} \ E^P\left[U_A(F(X_T)) - \int_0^T c_t(\nu_t)\,dt\right] \tag{2}$$

and the contracts such that the agent's objective value exceed the reservation utility R are defined to be the set

$$\Sigma = \{F(\cdot) | V_A(X_T) \geq R\} \tag{3}$$

Finally, the principal's problem is

$$V_P(X_T) = \sup_{\nu \in \Sigma} E\left[U_P(X_T - C_T)\right] \tag{4}$$

B. Robust Formulation

A preference for robustness is achieved by having an adverse alternative model with different probability measure Q, with

$$\frac{dQ}{dP}|_{\mathscr{F}_t} = e^{-\int_0^t \varphi_r dW_r - \frac{1}{2}\int_0^t \varphi_r^2 dr} \tag{5}$$

By Girsanov's theorem, the Q-Brownian motion satisfies,

$$dW_t^P = \varphi_t dt + dW_t^Q \tag{6}$$

, and the alternative model becomes

$$dX_t = \sigma(X, \beta_t)\left(\lambda_t(X, \alpha_t) + \varphi_t\right)dt + \sigma(X, \beta_t)\,dW_t^P \tag{7}$$

Consequently, the agent's and principal's objective function are changed to the max-min problem,

$$V_A(X_T) = \sup_{\nu} \inf_{\varphi} E^Q\Big[U_A(F(X_T)) - \int_0^T c_t(\nu_t)\,dt + \int_0^T \frac{\varphi_t^2}{2\Phi^A(X,t)}\,dt\Big] \tag{8}$$

,and

$$V_P(X_T) = \sup_{\nu \in \ \Sigma} E^Q\left[U_P(X_T - C_T)\right] \tag{9}$$

,where the agent's robustness preference is characterized by Φ^A. Observe that the extra term in the agent's objective function represents the entropy penalty when choosing the drift perturbation in the alternative model, which turns out to be the KL-divergence when $\Phi^A = 1$.

IV. Dynamic Programming Approach to Robust Problem

For tractbility, we will start by considering a less general model as proposed in [] **throughout this section** where

$$\sigma(X, \beta_t) = \beta_t; \quad \lambda_t(X, \alpha_t) = \lambda; \quad c_t(\nu_s) = 0$$

That is, the principal chooses both the contact and the effort of the agent and no moral hazard takes place. Despite the simplicity, this model serves as a perfect illustration for the tactics we proposed.

Observe that the absence of the cost of effort makes the problem solvable using convex duality method. Nevertheless, we shall continue to work on this setting to recover the result in [Cadenillas et al.(2007)Cadenillas, Cvitanić, and Zapatero] using our approach.

The agent and principal are respcetively maximizing

$$\inf_{\varphi} E^Q\Big[U_A(F(X_T)) + \int_0^T \frac{\varphi_t^2}{2\Phi^A(X,t)}\,dt\Big] \quad \text{and} \quad E^Q\left[U_P(X_T - C_T)\right] \tag{10}$$

A. The Agent's HJBI equation

In view of the agent's problem in (10), it is natural to consider the Hamilton-Jacobi-Bellman-Isaacs (HJBI) equations as suggested in the [Maenhout(2004)]. Assume the agent's value function $V(t,x)$ satisfies $V_{xx} < 0$, then the HJBI equation is readily written as

$$0 = \sup_{\beta} \inf_{\varphi}\left[v_t^A + (\lambda\beta_t + \beta_t\varphi_t)v_x^A + \frac{1}{2}\beta_t^2 v_{xx}^A + \frac{1}{2\Phi^A(x,t)}\varphi_t^2\right] \tag{11}$$

Solve for the minimization part of the problem yields

$$\varphi^* = -\Phi^A(x,t)\beta_t v_x^A \tag{12}$$

Substitute the result back to 11 gives

$$\sup_{\beta}\left[v_t^A + \lambda\beta_t v_x^A + \frac{1}{2}\beta_t^2 v_{xx}^A - \frac{1}{2}\Phi(x,t)^A\beta_t^2(v_x^A)^2\right] = 0 \tag{13}$$

Note that $\Phi^A(x,t)$ can be any $\mathbb{F}$-measureable function. However, for the problem at hand there is no anlytically tractable solution for a random Φ. To bridge over the gap and make the HJBI equation analytically tractable. **We propose that**

$$\Phi(x,t)^A = -\frac{c^A v_{xx}^A}{(v_x^A)^2} > 0 \quad \text{given some } c^A > 0 \tag{14}$$

The first-order condition implies the necessity condition for optimally is

$$\beta^* = -\frac{\lambda v_x^A}{(1+c^A)v_{xx}^A} \tag{15}$$

Therefore,

$$\varphi^* = -\frac{\lambda c^A}{1+c^A} \tag{16}$$

, and the HJBI equation becomes

$$v_t - \frac{1}{2}\frac{\lambda^2 (v_x^A)^2}{(1+c^A)v_{xx}^A} = 0 \tag{17}$$

Note that, if the agent desires no robustness ($c^A = 0$), then the problem reduces to the setting suggested in section 5.3 in [Cvitanić(2018)]. Moreover, the drift distortion is solely dependent on the constant c^A.

B. The Principal's HJB equation

Recall that in [Cvitanić(2018)], arthors show that the agent's value process admits a dynamic representation $Y_t^{Z,\Gamma}$. Given a pair of $\mathbb{F}$-predictable process (Z,Γ) with finite norm, the process $Y_t^{Z,\Gamma}$ is defined to be

$$\begin{aligned} dY_t^{Z,\Gamma} &= (Z_t\lambda\beta_t^* + \frac{1}{2}\Gamma_t(\beta_t^*)^2 - H_t)(Y_t^{Z,\Gamma}, Z_t, \Gamma_t)dt \\ &+ Z_t\beta_t^*(Y_t^{Z,\Gamma}, Z_t, \Gamma_t)dW_t^Q \end{aligned} \tag{18}$$

, where

$$\begin{aligned} H_t(Y_t^{Z,\Gamma}, Z_t, \Gamma_t) &= \sup_v h(Y_t^{Z,\Gamma}, Z_t, \Gamma_t, v) \\ &= \sup_v -c_t(v_t) + k_t(x,u)Y_t^{Z,\Gamma} \\ &+ (\lambda\beta_t + \beta_t\varphi_t)Z_t + \frac{1}{2}\beta_t^2\Gamma_t \\ &= \sup_v (\lambda\beta_t + \beta_t\varphi_t)Z_t + \frac{1}{2}\beta_t^2\Gamma_t \end{aligned} \tag{19}$$

is the Hamiltonian functional.

Proposition 1. *Given the drift distortion φ_t, if there exits a maximizer to the agent's Hamiltonian given, then under the contract with payoff $Y_T^{Z,\Gamma}$, the agent's value function coincides with the process $Y_t^{Z,\Gamma}$ and the corresponding agent's optimal actions are uniquely defined as the maximizer the .*

This proposition is a modification to the Proposition 3.3. in [Cvitanić(2018)] with the similar proof.

Together with Theorem 3.9 in[Cvitanić(2018)] , the proposition 1 allows us to write the HJB equation for the principal's value function with dependence over process $X_t, Y_t^{Z,\Gamma}$. For a given (z,γ), the optimal

$$\beta^* = -\frac{\lambda z}{(1+c^A)\gamma}$$

. With $w = \frac{z}{\gamma}$, the HJB equation then becomes,

$$\begin{cases} v_t + \dfrac{\lambda^2}{(1+c^A)^2} \sup\limits_{(z,w)\in\mathbb{R}^2} \Big[-v_x + \dfrac{c^A}{2} zwv_y + \\ \qquad \dfrac{1}{2}w^2(v_{xx} + z^2 v_{yy}) + w^2 z v_{xy} \Big] = 0 \\ v(T,x,y) = U_P(x - U_A^{-1}(y)) \end{cases} \tag{20}$$

The first-order conditions are

$$\begin{aligned} w^* &= \frac{2v_{yy}v_x + c^A v_{xy}v_y}{2v_{yy}v_{xx} - 2v_{xy}^2} \\ z^* &= -\frac{2v_{xy}v_x + c^A v_{xx}v_y}{2v_{yy}v_x + c^A v_{xy}v_y} \end{aligned} \tag{21}$$

We also have, by Ito's Lemma

$$\begin{aligned} dv_x &= -\frac{\lambda}{1+c^A} v_x dW_t^Q + D(v,x,y)dt \\ v_x(T,x,y) &= U_P'(x - U_A^(- 1)(y)) \end{aligned} \tag{22}$$

,where $D(v,x,y)$ is the drift function. Differentiating the HJB euqation for v with respect to x shows that the $D(v,x,y) = 0$ and we have

$$dv_x = -\frac{\lambda}{1+c^A} v_x dW_t^Q \tag{23}$$

, a standard stochastic differential equation with the solution

$$v_x(t, X_t, Y_t) = m_0 e^{-\frac{\lambda^2}{2(1+c^A)^2}t + \frac{\lambda}{1+c^A}W_t^Q} \tag{24}$$

Compared to the result in [Cadenillas et al.(2007)Cadenillas, Cvitanić, and Zapatero], introducing robustness simply scales the final differential equation by $1 + c^A$, a observation that is also pointed out in the proposition 2 in [Maenhout(2004)].

However, we have to admit that the setting in section IV is restrictive because enforcing the effort process to zero makes the problem less a standard moral hazard as we wish. Naturally, our next task is to generalize the previous result to the case with non-zero cost.

C. Generalization

For the general general case with non-zero effort process, a set of similar procedures can be applied to obtain the results. Sequentially, we may proceed by setting up the agent's HJBI equation, finding optimizer, then construct the principal's HJB equation based on Theorem 3.9 in [Cvitanić(2018)] and obtaining the solutions. However, the result may be a series of BSDE or SDE equations that are analytically intractable in general.

V. Conclusion

We generalized the dynamic approach proposed in [Cvitanić(2018)] to the robust prinicipal-agent problem with lump-sum payment, where the robustness preference takes place in the drift function. The method reformulates he agent's objective problem to be the max-min problem over a different measure Q and optimizes the principal's problem over a reduced family of contracts. Further work can be done to incorporate a general cost of effort, discount process and simulate the result.

ACKNOWLEDGMENT

I would like to acknowledge the funding support from Nanyang Technological University – URECA Undergraduate Research Programme for this research project.

REFERENCES

[Bebchuk and Fried(2004)] Lucian Arye Bebchuk and Jesse M Fried. Power and Pay. *Pay without performance: The unfulfilled promise of executive compensation*, (February):61–120, 2004.

[Cadenillas et al.(2007)Cadenillas, Cvitanić, and Zapatero] Abel Cadenillas, Jakša Cvitanić, and Fernando Zapatero. Optimal risk-sharing with effort and project choice. *Journal of Economic Theory*, 133(1):403–440, 2007. ISSN 00220531. doi: 10.1016/j.jet.2005.12.007.

[Cvitanić(2018)] Cvitanić. Dynamic programming approach to principal–agent problems. *Finance and Stochastics*, 22(1):1–37, 2018. ISSN 09492984. doi: 10.1007/s00780-017-0344-4.

[Gabrel et al.(2014)Gabrel, Murat, and Thiele] Virginie Gabrel, Cécile Murat, and Aurélie Thiele. Recent advances in robust optimization: An overview. *European Journal of Operational Research*, 235(3): 471–483, 2014. ISSN 03772217. doi: 10.1016/j.ejor.2013.09.036. URL http://dx.doi.org/10.1016/j.ejor.2013.09.036.

[Guo et al.(2019)Guo, Langrené, Loeper, and Ning] Ivan Guo, Nicolas Langrené, Grégoire Loeper, and Wei Ning. Robust utility maximization under model uncertainty via a penalization approach. 2019. URL http://arxiv.org/abs/1907.13345.

[Holmstrom and Milgrom(1987)] Bengt Holmstrom and Paul Milgrom. Aggregation and Linearity in the Provision of Intertemporal Incentives Author (s): Bengt Holmstrom and Paul Milgrom Published by : The Econometric Society Stable URL : http://www.jstor.org/stable/1913238 . AGGREGATION AND LINEARITY IN THE PROVISION OF I. *Econometrica*, 55(2):303–328, 1987. URL http://www.jstor.org/stable/1913238.

[Maenhout(2004)] Pascal J. Maenhout. Robust portfolio rules and asset pricing. *Review of Financial Studies*, 17(4):951–983, 2004. ISSN 08939454. doi: 10.1093/rfs/hhh003.

[Mirrlees(2017)] Author J A Mirrlees. Linked references are available on JSTOR for this article : The Theory of Moral Hazard and Unobservable Behaviour : Part I. 66(1):3–21, 2017.

[Patrick Bolton(2005)] Mathias Dewatripont Patrick Bolton. *Contract Theory*. 2005. ISBN 9781351712613. doi: 10.4324/9781315178189.

[Sannikov(2008)] Yuliy Sannikov. A continuous-time version of the principal-agent problem. *Review of Economic Studies*, 75(3):957–984, 2008. ISSN 00346527. doi: 10.1111/j.1467-937X.2008.00486.x.

[Zeckhauser(1970)] Richard Zeckhauser. Medical insurance: A case study of the tradeoff between risk spreading and appropriate incentives. *Journal of Economic Theory*, 2(1):10–26, 1970. ISSN 10957235. doi: 10.1016/0022-0531(70)90010-4.