Chapter 3
Time Inconsistency and Self-Control Optimization Problems: Progress and Challenges

Yun Shi and Xiangyu Cui

Abstract Time inconsistency has been an important issue in many stochastic decision problems arisen in real life and financial decision making, especially in the dynamic investment area. When a stochastic decision problem is time inconsistent, the decision maker would be puzzled by his/her conflicting decisions "optimally" derived from his/her time-varying preferences at different time instants. In the literature, the time inconsistent problem is also called the self-control problem, as the decision maker needs to exert proper self-control to resist present temptation and then achieve a better long-term performance. Different approaches dealing with time inconsistency in the literature are reviewed in this paper. After that, the open questions and challenges are also discussed.

Keywords Time inconsistency • Self-control • Present bias • Non-separable problem • Quasi-hyperbolic discounting function • Dynamic portfolio optimization

3.1 Introduction

Time inconsistency, also called dynamic inconsistency, refers to the phenomena that the long-term optimal decision policy determined at time 0 is no longer optimal when reconsidering the truncated decision problem at time t. Many real-life decision

Y. Shi
School of Management, Shanghai University, Shanghai, People's Republic of China
e-mail: y_shi@shu.edu.cn

X. Cui (✉)
School of Statistics and Management, Shanghai University of Finance and Economics, Shanghai, People's Republic of China
e-mail: cui.xiangyu@mail.shufe.edu.cn

problems are time inconsistent problems, such as whether to abstain from smoking.[1] Apparently, the long-term optimal decision is to insist abstaining from smoking. However, when coming into action, the smoker may prefer to smoke today and abstain tomorrow. The reason behind is that comparing to the present temptation (the pleasure of smoking), the decision maker evaluates the bad consequences received in the future (bad health) with a large discount factor, which is also called the present bias (see Thaler 1981; Loewenstein and Prelec 1992). In the literature, the time inconsistent problem is also called the self-control problem, as the decision maker needs to exert proper self-control in order to achieve a good long-term performance.

Essentially, when a decision maker faces a dynamic decision-making problem, his/her preferences at different time instants for the corresponding tail parts of the time horizon could be time-varying and/or state dependent. Actually, we all have faced occasions in which we change our minds, but usually we do not go to extraordinary steps to prevent ourselves from deviating from the original plan. The only circumstances in which we would want to commit ourselves to our planned course of action is when we have a good reason to believe that if we change our preferences later, this change of preferences will be a mistake. In the smoking case, the smoker definitely believes that abstaining from smoking is good for health. Thus, the crucial question is this: if I know I am going to change my mind about my preferences, when and how would I take some actions to restrict my future behaviour?

In this article, we start by summarizing some progress in dealing with the time inconsistency and self-control issue, especially in the dynamic investment area. We then describe some major open questions in this area.

3.2 Progress

3.2.1 Separable Problem Versus Non-separable Problem

The time inconsistent problems can be classified into two categories: separable ones and non-separable ones, which are caused by the present bias and the non-separable preference, respectively. Dynamic investment and consumption problems with quasi-hyperbolic discounting and dynamic mean-variance portfolio selection problems are two salient time inconsistent decision problems in the literature. The past few years have seen substantial progress in our understanding of the time consistency issue. Much of the progress concerns these two problems.

For the dynamic investment and consumption problem with quasi-hyperbolic discounting, the decision maker's time preference, which represents the value at time t of \$1 received at future time s, is described by

$$D_t(s) = \begin{cases} 1, & \text{if } s = t, \\ \beta\delta^{(s-t)}, & \text{if } s > t, \end{cases}$$

---

[1]More time inconsistency examples can be found in Gul and Pesendorfer (2001), Grenadier and Wang (2007), Björk and Murgoci (2010), Basak and Chabakauri (2010) and Cui et al. (2012).

where $0 < \beta < 1$ is the quasi-hyperbolic discounting parameter and represents the short-run discounting, and δ represents the long-run discounting (see Laibson 1997; O'Donoghue and Rabin 1999). The quasi-hyperbolic discounting is a typical and well-documented time preference with present bias, under which the decision maker tends to underestimate the value of payoff in the future. The decision maker's investment objective at time t is the expected sum of discounted utilities of the future consumptions and the terminal wealth,

$$\mathbb{E}_t \left[\sum_{s=t}^{T-1} D_t(s) U(c(s)) \right] + D_t(T) \mathbb{E}_t[U(X(T))].$$

Due to the existence of the short-run discounting parameter β, the preferences at different time instants are inconsistent. The decision maker may switch his/her mind at a later time t and prefer to consume more than his/her original plan to maximize his/her short-term preference (see Thaler and Shefrin 1981). The smaller the parameter β, the larger the conflict between the long-term optimal consumption plan and the short-term optimal consumption plan.

For dynamic mean-variance portfolio selection problem, the decision maker's preference at time t is the weighted sum of the conditional expected value and the conditional variance of the terminal wealth,

$$\mathbb{E}_t[X(T)] - \lambda \text{Var}_t(X(T)),$$

where $\lambda > 0$ is the risk aversion parameter. As the variance term does not satisfy the smooth property, i.e., $\text{Var}_t(X(T)) \neq \text{Var}_t(\text{Var}_s(X(T)))$ for $s > t$, the preferences at different time instants would be definitely inconsistent. After a mean-variance investor derives his/her long-term optimal investment strategy at time 0, he/she could be tempted to adopt a different strategy at a later time t in order to achieve a short-term mean-variance efficiency (see Basak and Chabakauri 2010; Cui et al. 2012).

One significant difference between the above two problems is that the objectives of the dynamic investment and consumption problem with quasi-hyperbolic discounting at different time instants only involve the separable expectation operator (which can be represented as the expected sum of the future discounted performance measures), while the objectives of the dynamic mean-variance portfolio selection problem at different time instants involve a non-separable operator: variance. In addition, almost all widely adopted risk measures in static portfolio selection, including variance, semi-variance, value-at-risk (VaR) and conditional value-at-risk (CVaR), become time inconsistent in dynamic mean-risk framework (see Boda and Filar 2006). Moreover, all those risk measures are of a non-separable nature.

3.2.2   Approaches Dealing with Time Inconsistency

Before summarizing the approaches, we need to know the mathematical meaning of time inconsistency. When a decision maker faces a time inconsistent dynamic decision problem, the overall objective for the entire time horizon under consideration does not conform with the local objective for a tail part of the entire time horizon. In the language of dynamic programming, *Bellman's principle of optimality* is not applicable in such situations, as the global and local interests derived from their respective objectives are not consistent (see Artzner et al. 2007).

Apparently, one direct approach to overcome the time inconsistency issue is to construct a time consistent decision model. This approach is widely studied in the field of dynamic risk measures and dynamic risk management. As a basic requirement, all the suitable dynamic risk measures should necessarily possess certain functional structure, such as

$$\rho_t(X_T) = \rho_t(-\rho_s(X_T)), \quad t < s < T,$$

in order to satisfy time consistency (see Rosazza Gianin 2006; Artzner et al. 2007; Jobert and Rogers 2008). When a dynamic risk measure is time consistent, it not only justifies the mathematical formulation for risk management, but also facilitates the solution process in finding the optimal decision (see Cherny 2010). However, this approach cannot be applied to the time inconsistent decision problems caused by present bias or non-separable preferences.

In the literature, there are mainly three different solution schemes in dealing with the time inconsistency issue for the general dynamic decision problem. The first solution scheme is the so-called pre-committed policy approach. In this approach, the decision maker strictly adheres to the global long-term optimal decision policy over the entire time horizon. In other words, the decision maker only cares about the global objective and fully ignores local objectives. Such policy is called the *pre-committed policy*.[2] To adopt the pre-committed policy, the decision maker can make the pre-committed policy the only feasible policy or the only economically reasonable policy via a strict self-control commitment or external contractual commitment, which is not easy in reality. Investment plan 401(k) is one such example in reality that forces employees not to withdraw pensions before retirement through a contractual penalty scheme (see Madrian and Shea 2001).[3]

---

[2]This policy is also termed "strategy of pre commitment" in Strotz (1956).

[3]In the United States, a 401(k) plan is the tax-qualified, defined-contribution pension account defined in section 401(k) of the Internal Revenue Code. The Internal Revenue Code imposes severe restrictions on withdrawals of pre-tax or Roth contributions while a person remains in service with the company and is under the age of 59.5. Any withdrawal that is permitted before the age of 59.5 is subject to an excise tax equal to ten percent of the amount distributed (on top of the ordinary income tax that has to be paid).

The second solution scheme is the time consistent policy approach. In this approach, the decision maker is aware of the inconsistency between the global objective and local objectives, but is unable to adhere to the pre-committed policy. Thus, the decision maker totally bows to local objectives (i.e., local temptations). As the decision maker at current time instant has a decision advantage with respect to the ones at future time instants, the decision maker's problem can be modelled as an intrapersonal sequential game. In the game, the decision maker at any time instant acts as a Stackelberg leader and makes his/her "best" decision by taking into account his/her decisions in future periods. The corresponding subgame perfect Nash equilibrium decision policy is called the *time consistent policy*.[4] This approach is widely applied to separable and non-separable time inconsistent decision problems. Laibson (1997), O'Donoghue and Rabin (1999, 2001), and Grenadier and Wang (2007) studied the time consistent policies for different financial decision problems with quasi-hyperbolic discounting. Basak and Chabakauri (2010), Hu et al. (2012), Lioui (2013), Chen et al. (2014b), Björk et al. (2014) and Cui et al. (2016) studied time consistent policies for the mean-variance preference under different market settings.

The third solution scheme is the self-control policy approach developed in the literature recently. In this approach, the decision maker intends to resolve conflicts between the long-term and short-term interests by reconciling the global objective and local objectives. To achieve this goal, the decision maker is required to possess a degree of willpower to exert self-control and resist the local temptation in the future time instants (see Rachlin 2004). Several theoretical models with a self-control feature have been developed to guide decision makers in achieving such a balance. For example, O'Donoghue and Rabin (1999, 2001) proposed the partial naive decision maker assumption, which assumes that the decision maker can exert self-control to have a larger quasi-hyperbolic discounting parameter, i.e.,

$$\hat{D}_t(s) = \begin{cases} 1, & \text{if } s = t, \\ \hat{\beta}\delta^{(s-t)}, & \text{if } s > t, \end{cases}$$

where $\beta < \hat{\beta} \leq 1$. With the larger quasi-hyperbolic discounting parameter $\hat{\beta}$, the decision maker decreases the conflict between the long-term and short-term preferences, and thus achieves a better balance between the long-term and short-term interests. Gul and Pesendorfer (2001, 2004) proposed the axiomatic theory of self-control, under which the decision maker integrates the opportunity costs of deviating from the local optimal policies into the global objective. By doing so, a policy taking into account both the long-term and short-term interests is obtained. Thaler and Shefrin (1981) and Bénabou and Pycia (2002) proposed the planner-doer model, and Fudenberg and Levine (2006, 2012) proposed the dual-self model,

---

[4]This policy is also termed the "strategy of consistent planning" in Strotz (1956), and the decision maker who adopts this policy is called the *sophisticated* decision maker in O'Donoghue and Rabin (1999, 2001) and Grenadier and Wang (2007).

which both assume that the global self can influence the myopic preferences of local selves through different self-control schemes and then derive the equilibrium policy between global self and local selves. We call this type of policies the *self-control policy* in this review.[5] These models with self-control features have been successfully applied to decision problems, whose time inconsistency is caused by present bias (see O'Donoghue and Rabin 1999, 2001; Grenadier and Wang 2007; Chen et al. 2014a; Tian 2016). However, these models are not applicable for non-separable time inconsistent decision problems.

Besides the above approaches, Cui et al. (2012) and Cui et al. (2015a) proposed a new angle for dealing with the time inconsistency of dynamic mean-variance portfolio problem. As mean-variance problem is a multi-objective optimization problem, a multi-objective version of principle of optimality is applied. In other words, any tail part of an efficient policy is also efficient for any realizable state at an intermediate period (see Li and Haimes 1987; Li 1990). In spirit of this logic, Cui et al. (2012) extended the concept of time consistency to a relaxed version to incorporate efficiency, namely, *time consistency in efficiency*. Through showing that the dynamic mean-variance formulation is not time consistent in efficiency, they demonstrated that the investor may have irrational local preferences of minimizing the risk and the return at the same time. Cui et al. (2012) relaxed the self-financing restriction and allow withdrawal of positive dollar amounts out of the market during the investment process. Furthermore, they proposed a better revised policy which can achieve the original mean-variance pair but obtain some extra (positive) dollar amounts with a strictly positive probability under certain probability distribution assumptions. Moreover, Cui et al. (2015a) studied the effect of portfolio constraints on the time consistency in efficiency of convex cone-constrained markets and established a general procedure for constructing time consistent in efficiency dynamic mean-variance portfolio selection problems by introducing suitable portfolio constraints.

3.3 Challenges

3.3.1 Dynamic Mean-Risk Portfolio Optimization Problems

Mean-risk portfolio selection models are widely used in portfolio management practices. Although most of these models suffered the problem of time inconsistency, only the dynamic mean-variance model attracts enough attentions in the literature.

---

[5]In Gul and Pesendorfer (2001, 2004), the preference of such a type of decision maker is called "preference with self-control". In O'Donoghue and Rabin (1999, 2001) and DellaVigna and Malmendier (2004, 2006), the decision maker who takes this type of policy is classified as *partially naive*. In Fudenberg and Levine (2006, 2012), this policy is termed "SR-Perfect equilibrium strategy".

Moreover, the published works on time inconsistency issue of dynamic mean-variance models mainly concentrate on the time consistent policy and the revised policy. Thus, there are several open questions in this field.

The first question is how to derive the time consistent policies for the mean-risk models beyond mean-variance. As these models are non-separable, there does not exist analytical or semi-analytical form policy. Thus, suitable numerical methods can be developed. Furthermore, the properties of the time consistent polices under different risk aversion parameter settings are worth investigation. Cui and Shi (2014) made an attempt to analyse the time consistent policy for multi-period mean-CVaR model with finite states.

The second question is whether the mean-risk models beyond mean-variance satisfy time consistency in efficiency. If not, the decision maker may devote himself/herself to constructing the revised policies, which is better than the pre-committed policy.

The third question is how to construct the self-control policies for the mean-risk models. Although the existing theoretical models with a self-control feature are not directly applicable to the non-separable mean-risk models, the idea is very useful in constructing some new theoretical models under non-separable mean-risk framework. To our best knowledge, there are some preliminary work in this direction. By extending the planner-doer model of Thaler and Shefrin (1981) and Bénabou and Pycia (2002), Cui et al. (2017) developed a two-tier planner-doer game framework with self-coordination, which is theoretically applicable to discrete-time non-separable decision problems. They successfully applied the proposed framework to deal with dynamic mean-variance portfolio selection problem and a two-period mean-CVaR portfolio selection problem. Similarly, Cui et al. (2015b) extended the dual-self model of Fudenberg and Levine (2006, 2012) and proposed a two-tier dual-self game model, which is theoretically applicable to continuous-time non-separable decision problems. Although the above two new frameworks have an important theoretical value, how to apply them to construct suitable self-coordination schemes and compute the corresponding self-control policies for the mean-risk models beyond mean-variance are still unclear.

3.3.2 Time Inconsistency Generated by Probability Weighting

The probability weighting function, proposed by Tversky and Kahneman (1992), transforms objective probabilities into decision weights. The original motivation for this transformation function was the simultaneous demand many people had for both lotteries and insurance. Typically, people prefer a 0.001 chance of $50,000 to a certain $50 but meanwhile prefer to pay $50 rather than face a 0.001 chance of a $50,000 loss. This combination of behaviours is difficult to explain under the expected utility theory. However, under the probability weighting framework, the unlikely events—gaining or losing $50,000—are overweighted, thereby explaining these choices. The probability weighting is a key feature of many behaviour

portfolio selection models, such as the rank-dependent utility model (see Schmeidler 1989; Abdellaoui 2002) and the cumulative prospect theory model (see Tversky and Kahneman 1992; He and Zhou 2011).

In a dynamic setting, probability weighting also generates time inconsistency (see Barberis 2012), which, once again, lies in the domain of non-separability. This inconsistency may be useful for understanding some real trading behaviours, for example, people sometimes hold on to losing investments longer than they were planning to, known as the disposition effect in the literature (see Odean 1998). However, there is relatively little research on it, especially when compared to the large literature on the inconsistency generated by present bias. Shi et al. (2015) suggested one possible approach to analyze the time inconsistency generated by probability weighting in dynamic setting, but other approaches are surely also possible and deserved more studies.

3.3.3 Data Challenge

As long as the data becomes more and more easy to collect, the decision makers begin to formulate their decision problems based on rich data. On the one hand, they can use the rich data to describe the dynamics of the uncertainties, which makes the constructions of dynamic decision problems possible. On the other hand, they may build data-driven decision problems by fully using the rich data (see Bertsimas and Thiele 2006; Delage and Ye 2010; Hou and Wang 2013; Huh et al. 2011 for data-driven decision-making examples). Based on these two developing directions, there will be more and more data-driven dynamic decision problems in research and practices.

In general, these data-driven dynamic decision problems are time inconsistent. Comparing to the dynamic decision problems with explicit assumptions on the uncertainties, the data-driven dynamic decision problems may introduce great computation challenges.

References

M. Abdellaoui, A genuine rank-dependent generalization of the Von Neumann-Morgenstern expected utility theorem. Econometrica **70**, 717–736 (2002)

P. Artzner, F. Delbaen, J.M. Eber, D. Heath, H. Ku, Coherent multiperiod risk adjusted values and Bellman's principle. Ann. Oper. Res. **152**, 5–22 (2007)

S. Basak, G. Chabakauri, Dynamic mean-variance asset allocation. Rev. Financ. Stud. **23**, 2970–3016 (2010)

N. Barberis, A model of casino gambling. Manag. Sci. **58**(1), 35–51 (2012)

D. Bertsimas, A. Thiele, Robust and data-driven optimization: modern decision-making under uncertainty. Tutor. Oper. Res. **4**, 95–122 (2006)

R. Bénabou, M. Pycia, Dynamic inconsistency and self-control: a planner-doer interpretation. Econ. Lett. **77**, 419–424 (2002)

T. Björk, A. Murgoci, A general theory of Markovian time inconsistent stochasitc control problem, working paper (2010). Available at SSRN: http://ssrn.com/abstract=1694759

T. Björk, A. Murgoci, X.Y. Zhou, Mean-variance portfolio optimization with state dependent risk aversion. Math. Financ. **24**, 1–24 (2014)

K. Boda, J.A. Filar, Time consistent dynamic risk measures. Math. Methods Oper. Res. **63**, 169–186 (2006)

S.M. Chen, Z.F. Li, Y. Zeng, Optimal dividend strategies with time-inconsistent preferences. J. Econ. Dyn. Control. **46**, 150–172 (2014a)

Z.P. Chen, G. Li, Y.G. Zhao, Time-consistent investment policies in Markovian markets: a case of mean-variance analysis. J. Econ. Dyn. Control **40**, 293–316 (2014b)

A.S. Cherny, Risk-reward optimization with discrete-time conherent risk. Math. Financ. **20**, 571–595 (2010)

X.Y. Cui, Y. Shi, Multiperiod mean-CVaR portfolio selection. in *Modelling, Computation and Optimization in Information Systems and Management Sciences*. (Springer, Berlin, 2014), pp. 293–304

X.Y. Cui, D. Li, S.Y. Wang, S.S. Zhu, Better than dynamic mean-variance: time inconsistency and free cash flow stream. Math. Financ. **22**, 346–378 (2012)

X.Y. Cui, D. Li, X. Li, Mean-variance policy for discrete-time cone-constrained markets: time consistency in efficiency and the minimum-variance signed supermartingale measure. Math. Financ. (2015a) doi:10.1111/mafi.12093

X.Y. Cui, X. Li, Y. Shi, Resolving time inconsistency of decision problem with non-expectation operator: from internal conflict to internal harmony by strategy of self-coordination. Working Paper (2015b)

X.Y. Cui, D. Li, Y. Shi, Self-coordination in time inconsistent stochastic decision problems: A planner-doer game framework. J. Econ. Dyn. Control **75**, 91–113 (2017)

X.Y. Cui, L. Xu, Y. Zeng, Continuous time mean-variance portfolio optimization with piecewise state-dependent risk aversion. Optim. Lett. **10**, 1681–1691 (2016)

E. Delage, Y.Y. Ye, Distributionally robust optimization under moment uncertainty with application to data-driven problems. Oper. Res. **58**, 595–612 (2010)

S. DellaVigna, U. Malmendier, Contract design and self-control: theory and evidence. Q. J. Econ. **119**, 353–402 (2004)

S. DellaVigna, U. Malmendier, Overestimating self-control: evidence from the health club industry. Am. Econ. Rev. **96**, 694–19 (2006)

D. Fudenberg, D.K. Levine, A dual-self model of impulse control. Am. Econ. Rev. **96**, 1449–1476 (2006)

D. Fudenberg, D.K. Levine, Timing and self-control. Econometrica **80**, 1–42 (2012)

S.R. Grenadier, N. Wang, Investment under uncertainty and time-inconsistent preferences. J. Financ. Econ. **84**, 2–39 (2007)

F. Gul, W. Pesendorfer, Temptation and self-control. Econometrica **69**, 1403–1435 (2001)

F. Gul, W. Pesendorfer, Self-control and the theory of consumption. Econometrica **71**, 119–158 (2004)

X.D. He, X.Y. Zhou, Portfolio choice under cumulative prospect theory: an analytical treatment. Manag. Sci. **57**, 315–331 (2011)

Z.S. Hou, Z. Wang, From model-based control to data-driven control: survey, classification and perspective. Inf. Sci. **235**, 3–35 (2012)

Y. Hu, H. Jin, X. Zhou, Time-inconsistent stochastic linear-quadratic control. SIAM J. Control Optim. **50**, 1548–1572 (2012)

W.T. Huh, R. Levi, P. Rusmevichientong, J.B. Orlin, Adaptive data-driven inventory control with censored demand based on Kaplan-Meier estimator. Oper. Res. **59**, 929–941 (2011)

A. Jobert, L.C. Rogers, Valuations and dynamic convex risk measures. Math. Financ. **18**, 1–22 (2008)

D. Laibson, Golden eggs and hyperbolic discounting. Q. J. Econ. **112**, 443–477 (1997)

D. Li, Multiple objectives and nonseparability in stochastic dynamic programming. Int. J. Syst. Sci. **21**, 933–950 (1990)

D. Li, Y.Y. Haimes, The envelope approach for multiobjective optimization problems. IEEE Trans. Syst. Man Cybern. **17**, 1026–1038 (1987)

A. Lioui, Time consistent vs. time inconsistent dynamic asset allocation: some utility cost calculations for mean variance preferences. J. Econ. Dyn. Control. **37**, 1066–1096 (2013)

G. Loewenstein, D. Prelec, Anomalies in intertemporal choice: evidence and an interpretation. Q. J. Econ. **107**, 573–598 (1992)

B.C. Madrian, D. Shea, The power of suggestion: inertia in 401(k) participation and savings behavior. Q. J. Econ. **116**, 1149–1187 (2001)

T. Odean, Are investors reluctant to realize their losses? J. Financ. **53**(5), 1775–1798 (1998)

T. O'Donoghue, M. Rabin, Doing it now or later. Am. Econ. Rev. **89**, 103–124 (1999)

T. O'Donoghue, M. Rabin, Choice and procrastination. Q. J. Econ. **116**, 121–160 (2001)

H. Rachlin, *The Science of Self-Control* (Harvard University Press, Cambridge, 2004)

E. Rosazza Gianin, Risk measures via g-expectations. Insur. Math. Econ. **39**, 19–34 (2006)

D. Schmeidler, Subject probability and expected utility without additivity. Econometrica **57**, 571–587 (1989)

Y. Shi, X.Y. Cui, D. Li, Discrete-time behavioral portfolio selection under cumulative prospect theory. J. Econ. Dyn. Control **61**, 283–302 (2015)

R.H. Strotz, Myopia and inconsistency in dynamic utility maximization. Rev. Econ. Stud. **23**, 165–180 (1956)

R.H. Thaler, Some empirical evidence on dynamic inconsistency. Econ. Lett. **8**, 201–207 (1981)

R.H. Thaler, H.M. Shefrin, An economic theory of self-control. J. Polit. Econ. **89**, 392–406 (1981)

Y. Tian, Optimal capital structure and investment decisions under time-inconsistent preferences. J. Econ. Dyn. Control **65**, 83–104 (2016)

A. Tversky, D. Kahneman, Advances in prospect theory: cumulative representation of uncertainty. J. Risk Uncertain. **5**(4), 297–323 (1992)