The recursive approach to time inconsistency

Andrew Caplin, John Leahy*

Department of Economics, New York University, New York, NY 10003, USA

Received 11 February 2004; final version received 24 May 2005
Available online 21 September 2005

Abstract

We introduce a new class of finite horizon stochastic decision problems in which preferences change over time, and provide a proof of the existence of a recursively optimal strategy. Recursive optimization techniques dominate many areas of economic dynamics. However, in decision problems in which tastes change over time, the solution technique most commonly applied is not recursive, but rather strategic (subgame perfection). In this paper we argue in favor of the recursive approach, and we take the necessary theoretical steps to make the recursive methodology applicable.

JEL classification: D80; D90

Keywords: Time consistency; Consistent plans; Subgame perfection

1. Introduction

Recursive optimization techniques dominate many areas of economic dynamics. However, in decision problems in which tastes change over time, the solution technique most commonly applied is not recursive, but rather strategic (subgame perfection). In this paper we argue in favor of a recursive approach to time inconsistency rather than the strategic approach, and we take the necessary theoretical steps to make the recursive methodology applicable. We introduce a new general class of finite horizon stochastic decision problems with time inconsistent preferences, and prove the existence of a recursively optimal strategy.

* Corresponding author. Department of Economics, New York University, 267 Mercer Street, 7th Floor, New York, NY 10003, USA.

E-mail address: john.leahy@nyu.edu (J. Leahy).

0022-0531/$ - see front matter © 2005 Elsevier Inc. All rights reserved.
doi:10.1016/j.jet.2005.05.006
The class of decision problems that we analyze has two features that distinguish it from the classical model of dynamic programming (c.f. [26]). The first critical difference is that we allow for a quite general form of time inconsistency. Not only do payoffs change over time, but we generalize the payoff function so that the entire probability distribution over future states and choices is allowed to impact payoffs in a non-linear fashion. This means that we include in our model situations in which current payoffs are qualitatively more complex than in the current theory, in that they may depend directly on beliefs about the future. The second difference is that we add some additional sources of uncertainty that are not included in the standard model. This change is needed for technical reasons, since it turns out that the existence of optimal strategies hinges on smoothing out potential discontinuities in the value function. Overall, we believe that our conditions for existence of recursively optimal strategies are applicable to almost all existing finite horizon models with time inconsistency.

In a classic paper, Strotz [27] introduced the concept of time inconsistent preferences, and discussed the impact that these preferences may have on consumption and saving. Since then, time inconsistency has been shown to arise from a number of sources including non-exponential discounting (Strotz), anticipatory emotions such as fear and anxiety [20,7], turnover in organizations [6], and commitment problems in games. It has been applied to many areas of economics and public policy including saving [19], monetary policy [18], capital taxation [12], crime [24], and procrastination [22].

In addition to introducing the model of changing tastes, Strotz outlined several different methods for solving the model, including what has become known as “sophisticated choice.” According to Strotz, the sophisticated decision maker is aware that his preferences may change in the future, and realizes that this may make it impossible for him to carry out certain plans that maximize utility from the current perspective. The decision maker is forced to

“modify his chosen plan to take account of future disobedience, realizing that the possibility of disobedience imposes a further constraint—beyond the budget constraint—on the set of plans which are attainable.”

Strotz suggested a recursive approach to solving such models. In the last period, since there is no longer any room for tastes to change in the future, the agent simply optimizes subject to the physical constraints. In each earlier period, the agent chooses “the best plan among those he will actually follow.” In other words, the agent maximizes subject to the constraint that future behavior is optimal. Note the strong analogy between the suggestion of Strotz and the recursive approach to decision making embodied in the theory of dynamic programming. In particular, the feasible set of options is generated recursively, via backward elimination of strictly dominated strategies.

Rather than apply the recursive methodology of Strotz, the current literature views the problem from an altogether different perspective. Following Peleg and Yaari [23], the sophisticated decision maker is viewed as a collection of distinct players, each making a choice at a different time based on the preferences prevailing at that time.¹ This intrapersonal game

¹ See Laibson [19] and O’Donohue and Rabin [22] for recent examples of this approach.
is then solved using the same solution concepts that would be applied in the corresponding multi-player game. Typically, the solution to the game is taken to be the set of subgame perfect equilibria.

One reason for the abandonment of the recursive approach was the recognition that there are settings in which no recursive optimum exists. Peleg and Yaari provided an example intended to demonstrate this non-existence, and while their example was incomplete (as noted by Bernheim and Ray [6]), they were essentially correct (see Section 2). With three or more periods, there may be no recursively optimal strategy. Given the non-existence of a recursive optimum, the appeal of the subgame perfect equilibrium can be understood in part as a desire to find solutions to this important class of models.

In this paper we return to the recursive approach proposed by Strotz. We consider a general finite-horizon problem with changing tastes and exogenous shocks. We provide examples that illustrate the difference between the recursive approach and the strategic approach. The examples enable us not only to pinpoint our reasons for favoring the recursive approach, but also to illustrate the factors influencing the existence of a recursive optimum. We use our example of non-existence to motivate the search for sufficient conditions for the existence of a recursive optimum. We are able to prove the existence of a recursive optimum in the class of models in which the future endogenous state is a noisy function of the current state. This class of models includes consumption-savings problems with random returns, models in which agents make mistakes, inventory models with random production or sales, models in which agents imperfectly observe their state, and search models.

In presenting our conditions for existence of an optimal strategy, we first provide appropriate modifications to the standard theory of dynamic programming. The main change that we make is to the objective function. In the standard theory, the objective is to maximize the expected present value of period payoffs, which, in turn, are functions of the current state variables and choices. In contrast, we allow the entire distribution of future states and choices to affect payoffs. Probabilities enter payoffs directly, and not only as expectational weights. This allows us to model situations in which agents respond to future uncertainty with feelings of fear, suspense or anxiety [7]. It also allows us to model situations in which the agent anticipates experiencing elation or disappointment if outcomes differ from expectations [21], and forms the decision theoretic basis for the study of games in which payoffs depend on beliefs [14]. The case in which period payoffs depend only on states and choices arises as a special case of our model.

In addition to providing positive results on the recursive approach, the other main thrust of the paper is to clarify the distinction between the recursive approach and the various alternative approaches to time inconsistency. We discuss the relationship of the recursive solution not only to the subgame perfect equilibria, but also to various existing refinements of perfection, including coalition-proofness [5].

The paper is organized as follows. The next section presents three examples that illustrate the differences between the recursive and game theoretic approaches to time inconsistency, and that illustrate the possibility of the non-existence of the recursive solution. Sections 3 and 4 develop general model of decision making with time changing preferences, and

---

2 The theorem of Harris [15] establishes existence in the two-period problem (see [7]).
show that an recursive solution exists in the case in which the future endogenous state variable is imperfectly related to choices today, a situation that naturally arises in many decision problems especially consumption-savings problems with random returns. Section 5 provides some commentary on the technical aspects of our approach, and relates our work to prior results. Section 6 relates our recursive model to the strategic approach. Section 7 concludes.

2. Perfection or recursion?

The standard approach to time inconsistency is to solve for the subgame perfect equilibria of the intrapersonal game played between the various temporal selves. We provide examples that pinpoint the difference between this approach and the recursive approach of Strotz. In addition to highlighting points of difference, the third example demonstrates the possibility of non-existence of the recursive solution. We close the section by outlining our approach to resolving the existence problem.

2.1. Example 1: a matter of indifference

Consider the two-stage decision problem illustrated in Fig. 1. In the first period the agent chooses between actions $A$ and $B$. If the first period choice is $A$, then in the second period the agent chooses between $a$ and $b$, while if the first period choice is $B$, then the subsequent choice is between $c$ and $d$. The payoffs to each combination of actions are given at each terminal node, with the payoff from the perspective of the first period given first.

The key feature of the example is the second period indifference between choice $a$ and choice $b$. The initial choice depends on the individual’s belief as to how he will break this indifference in period 2. In fact, it is clear that there are two pure strategy subgame perfect equilibria, and that these are $\{A, a, c\}$ and $\{B, b, c\}$. The equilibrium $\{A, a, c\}$ is motivated by the initial self’s certainty that future self will pick $a$ in response to $A$. The equilibrium $\{B, b, c\}$ is motivated by the conviction that future choice will be $b$.

In contrast to the multiplicity of equilibria with the strategic approach, there is only one recursive optimum. The recursive approach tells the agent to choose the best plan that will actually be followed. In this case there are three plans that satisfy the requirement that they are acceptable to the agent in period 2. Specifically, $\{A, a\}$, $\{A, b\}$ and $\{B, c\}$ are all consistent with optimal behavior in the second period. The remaining plan, $\{B, d\}$, will be rejected in favor of $\{B, c\}$ in the second period. Among the consistent plans $\{A, a\}$ is clearly the best choice from the period 1 perspective, and is therefore the recursive solution.

As the example suggests, the distinction between subgame perfection and recursion comes down to distinct visions of how the agent believes future selves will select from among the policies that are then optimal. With subgame perfection, the agent has definite but arbitrary beliefs about the breaking of indifference. In contrast, the recursive approach works by excluding plans that are suboptimal for the future self, and then leaving the breaking of indifference to the initial self. In effect the agent makes two decisions. The agent chooses an action today and an optimal continuation plan from the set of optimal continuation plans.
In the example both $a$ and $b$ are optimal second period responses to $A$; preferences in the first period break this indifference.\(^3\)

On intuitive grounds, we find the vision in which the current self gets to break indifference among future strategies more appealing than the vision of a well-defined but arbitrary belief-selection mechanism. In so far as the future self exists in the current period, it exists only in the mind of the current self. It is not easy to understand how the current self would convince itself that the future self was going to make a selection that would damage the current self, as is called for in the perfect equilibrium $\{B, b, c\}$.

2.2. Example 2: commitment and indifference

One possible response to the example above is to claim that indifference shows up only rarely, so that the issue of how choices get resolved when there is indifference is unimportant. Our next example shows that indifference arises naturally in settings with time inconsistency.

There are two periods. In the second period an agent chooses between two alternatives, $y_2 \in \{A, B\}$. The time inconsistency arises because $A$ is preferred to $B$ in the second period,

\(^3\) Several recent papers employ similar logic to select among subgame perfect equilibria. Diamond and Köszegi [11] in a model of retirement with hyperbolic discounting also assume that the indifference of later selves may be broken by the strict preference of earlier selves. Benabou and Tirole [3] note “when multiple PBE’s exist in our model we shall generally emphasize the Pareto-dominating one(s).”
whereas $B$ is preferred to $A$ in the first period. In the first period, the agent makes a costly decision $x \geq 0$ that affects the desirability of the alternatives in the second period; the higher is $x$, the greater is the second period incentive to choose $B$, but the greater is the cost in the first period. One interpretation is that $x$ represents a technology that allows the agent to commit to choose $B$.

The optimal choice of $x$ in this example is clear. Since commitment is costly, the agent should choose the level of $x$ that just tips the balance in favor of $B$. The way that the strategic approach and the recursive approach arrive at this conclusion, however, is revealing. We will therefore proceed with the formal analysis.

Let the payoff in period 2 be

$$U_2(x, A) = 1;$$
$$U_2(x, B) = x.$$

Note that the agent prefers $A$ when $x = 0$, but that higher values of $x$ increase the desirability of $B$. The payoff in period 1 depends on the choice of $x$ and the expected decision in the second period:

$$U_1(x) = -\lambda x + E_1 I_{\{y_2 = B | x\}}.$$

Here $I_Z$ denotes the indicator of the event $Z$, $E_t$ represents the mathematical expectation conditional on period $t$ information, and $\lambda$ reflects the costliness of increasing $x$. We assume that $\lambda < \frac{1}{2}$.

We arrive at the subgame perfect equilibrium through backward induction. In period 2, the optimal choice is $A$ if $x$ is strictly below 1, and $B$ if $x$ is strictly above 1. If $x = 1$, the agent is indifferent. In this case, any mixed strategy on the part of the period 2 self can form part of a subgame perfect equilibrium. Let $\pi \in [0, 1]$ denote the probability that the period 2 self will choose action $B$ when $x = 1$.

One subgame perfect equilibrium involves the choice of $x = 1$ in period 1, together with the certainty on the part of the period 1 self that the period 2 self will select action $B$ in period 2; that is $\pi = 1$. The payoff to the period 1 self from this strategy is $1 - \lambda > 0$, and the payoff to the period 2 self is 1. The fact that this is an equilibrium follows from the fact that higher choices of $x$ only reduce the period 1 payoff, whereas choices of $x < 1$ result in a period one payoff that is negative. The agent in period 2 is indifferent between $A$ and $B$ and therefore willingly chooses $B$.

Similar logic explains why there is no other subgame perfect equilibrium. Suppose that period 1 self believes that $\pi < 1$. In this case the period 1 payoff is $\pi - \lambda$. The period 1 self can get a strictly higher payoff by raising $x$ by $\varepsilon > 0$ and ensuring that the second period choice is $B$. In this case, no equilibrium exists since the first period payoff is strictly decreasing in $\varepsilon$ and fails to achieve its limit value of $1 - \lambda$.

In contrast to this somewhat intricate logic, the reasoning behind the recursive approach is straightforward. The agent in period 1 selects $x = 1$. This makes the period 2 agent indifferent between the two choices, in which case the recursion hands the decision to the agent in period 1, who decides that $B$ is to be the period 2 choice in the face of indifference. Period 2 self has no objections, and the strategy proceeds.
We believe that one disturbing aspect of the perfect equilibrium logic is that it pins down second period choice in the face of indifference not for behavioral reasons, but for technical reasons based on the need to guarantee existence. Only $\pi = 1$ is consistent with existence. In contrast, the recursive approach pins down second period choices for behavioral reasons. Either the second-period self prefers the action, or, if the second-period self is indifferent, the first-period self prefers the action.

The models of Battaglini, Benabou, and Tirole [2] and Benabou and Tirole [3] share some similarities with the above example. In each case indifference is generic and is due to a desire of the period-one self to push the period-two self over a threshold. In these models a weak-willed period-one self trades off the cost of persevering with a difficult task today and the benefit of inducing the period-two self to take a desirable action tomorrow. The optimal policy is to persevere just enough to induce the period-two self to take the desirable action. It is interesting to note that Benabou and Tirole propose a game-theoretic solution that differs from the recursive solution described here. Their solution involves both selves playing mixed strategies such that each self is indifferent between the various actions at their disposal. In contrast, the recursive approach would have the period-one self choosing a probability of persevering that makes the period-two self indifferent between his various choices tomorrow; the period-one self would then break this indifference in the direction he desired.

2.3. Example 3: non-existence of a recursive optimum

An example of non-existence of a recursive optimum is readily generated by slightly amending Example 2. We add a third period between periods 1 and 2, and label it period $1^\ast$. The innovation is that the period $1^\ast$ self not only prefers $A$ to $B$ as does the period 2 self, but does so to an even greater extent than does the period 2 self:

$$U_{1^\ast} = 2E_{1^\ast}I_{[y_2 = A|x]}.$$ 

All else is the same as Example 2.

The introduction of the extra period does not alter the perfect equilibria of the game. The only choices are the period 1 choice of $x$, and the period 2 choice between $A$ and $B$. The only equilibrium involves $x = 1$ and the subjective certainty on the part of the period 1 self that the period 2 self will select $B$ in the face of the ultimate indifference. Nothing changes since the period $1^\ast$ self has no choices to make.

The addition of the extra period, however, does affect the recursive solution. Here the period $1^\ast$ agent may break the period 2 indifference between $A$ and $B$ in favor of $A$. The agent in period 1 will anticipate this and choose $x$ slightly greater than 1, in order to ensure that the second period self favors $B$. Herein lies the existence problem. The optimal strategy is to break indifference in the smallest way possible, which is ill-defined.

From the recursive viewpoint, the two examples really are different, and one can provide a reasonable “cheap talk” reason for the difference. In the two-period example, the period 1 self could write a note to the period 2 self saying “when indifferent, please do what I would prefer, which is action $B$.” The period 2 self would receive this note, and have no reason to deviate. When trying a similar exercise in the current example, the period 1 self would write exactly the same note. Unfortunately, the period $1^\ast$ self would rip the note up, and hand
on a note with the alternative instruction “when indifferent, please do what I would prefer, which is action A.” The period 2 self would again have no reason to deviate. Anticipating this reaction, the period 1 self chooses a slightly higher value for x.

While the example shows non-existence to be an issue, we do not feel that it demonstrates a conceptual flaw in the recursive approach. We do not discard the principle of maximization because there is no maximal element in the open interval (0, 1). Neither should we discard recursive optimization because of the non-existence of a solution to a more sophisticated optimization problem.

2.4. Toward a proof of existence

Given our preference for the recursive approach, the natural next step is to search for acceptably broad conditions that guarantee existence of recursively optimal strategies. In order to identify such conditions, it is important to understand the roots of the existence problem in the third example above. To do this, we use the traditional language of dynamic programming to describe the problem.

Let \( V_2(x) \) denote the maximized payoff in the second period as a function of the prior choice of \( x \). The second period value function is continuous as a function of the past choices,

\[
V_2(x) = \begin{cases} 
  x & \text{if } x \geq 1; \\
  1 & \text{if } x \leq 1.
\end{cases}
\]

The optimal choice correspondence is upper hemi-continuous but not continuous,

\[
G_2(x) = \begin{cases} 
  B & \text{if } x > 1; \\
  A \cup B & \text{if } x = 1; \\
  A & \text{if } x < 1.
\end{cases}
\]

In period 1\(^*\), the only choice that the agent makes is to break indifference in the second period. The second period choice correspondence is therefore a constraint on the period-1\(^*\) problem. This correspondence is not continuous, so the standard theorem of the maximum no longer applies. It turns out that the period 1\(^*\) value function is upper hemi-continuous with a discontinuity at \( x = 1 \),\(^4\)

\[
V_{1\,*}(x) = \begin{cases} 
  0 & \text{if } x > 1; \\
  2 & \text{if } x \leq 1;
\end{cases}
\]

and that the period 1\(^*\) choice correspondence is no longer upper hemi-continuous in \( x \):

\[
G_{1\,*}(x) = \begin{cases} 
  B & \text{if } x > 1; \\
  A & \text{if } x \leq 1.
\end{cases}
\]

\(^4\) It is easy to show that this is a general result. If the constraint correspondence is upper hemi-continuous, then under the remaining conditions of the theorem of the maximum, the value function is upper semi-continuous.
Note that the interpretation of this is not that the choice between $A$ and $B$ is actually made in period 1*, but rather that the period 1* self ensures that period 2 self breaks indifference by picking $A$ when $x = 1$.

The existence problem arises in period 1, when the value function fails to achieve its supremum of $1 - \lambda$,

$$V_1 = \max \left\{ \sup_{x \in [0,1]} -\lambda x, \sup_{x > 1} 1 - \lambda x \right\} = 1 - \lambda.$$

The problem is that the period 1* optimal choice correspondence, $G_1^*$, enters as a constraint in period 1, and this correspondence is not upper hemi-continuous.

In the next section we present sufficient conditions for the existence of recursively optimal strategies. We consider environments in which the endogenous state tomorrow is not a determinate function of the choice today. This uncertainty will “smooth away” the effect of discontinuities in payoffs in subsequent periods. Situations in which such noise is natural include any problem in which the agent cannot perfectly control the endogenous state variable.

3. The general model

3.1. The decision making environment

We begin by summarizing the decision making environment. Time is discrete and there is a finite horizon, $T \geq 0$. For each $0 \leq t \leq T$, let $X_t$ and $Z_t$ be compact subsets of finite-dimensional Euclidean spaces. Let $S_t = X_t \times Z_t$. $X_t$ is the set of endogenous state variables in period $t$; $Z_t$ is the set of exogenous state variables; and $S_t$ is the overall state of the system. An initial condition $s_0 \in S_0$ is given.

In all periods, the agent observes the state and then makes a decision. The action space for period $t$ is $Y_t$ which we assume to be a compact subset of a Euclidean space. The set of actions available to the agent in period $t$ depends on the current state of the system and is given by the action correspondence $\Gamma_t : X_t \times Z_t \rightarrow Y_t$. We assume that $\Gamma_t$ satisfies the standard regularity conditions.

Assumption 1. $X_t$, $Y_t$ and $Z_t$ are compact subsets of finite-dimensional Euclidean spaces. $\Gamma_t(x_t, z_t)$ is a measurable correspondence, which, given $z_t \in Z_t$, is non-empty, compact-valued, and continuous in $x_t$.

3.2. The laws of motion

We make the following assumption regarding the evolution of the system:

Assumption 2. The evolution of the system is represented by a transition probability kernel $\nu_{y_t, z_t}(dx_{t+1}, dz_{t+1}) = \lambda_t(y_t, z_t, x_{t+1}) dx_{t+1} q_t(z_t, dz_{t+1})$ such that $\lambda_t$ is bounded and continuous in all of its arguments and $q_t$ is simple.
Assumption 2 is where we introduce noise into the decision problem. The endogenous state variable $x_{t+1}$ is random and its distribution conditional on $s_t$ is absolutely continuous with respect to the Borel measure on $S_{t+1}$.\(^5\)

The timing is conventional. The agent observes the initial state $s_0$ and then selects an action from $\Gamma_0(x_0, z_0)$. At the very start of period 1, $s_1$ is realized, whereupon the cycle then repeats itself in the obvious manner.

### 3.3. Payoffs

We now turn to the computation of payoffs, where the issue of time inconsistency appears. In order to encompass models in which beliefs enter utility directly (see the discussion in Section 6), we consider utility over temporal lotteries as in Kreps and Porteus [17].

We begin with some notation. Given an abstract space $\Theta$, let $P(\Theta)$ denote the space of Borel probability measures over $\Theta$ endowed with the topology of weak convergence. Note that if $\Theta$ is a compact, separable metric space, then $P(\Theta)$ is compact, separable and metrizable (see [17]).

Define $P_{T+1} = \{1\}$. $P_{T+1}$ plays no role in the analysis other than to allow the definition of the pair $(y_T, p_{T+1}) \in Y_T \times P_{T+1}$ and thereby simplify the notation. For $t \leq T$, define $P_t = P(S_t \times Y_t \times P_{t+1})$. The $P_t$ describe in a recursive manner probability distributions over future states and actions. The recursivity will prove useful when discussing optimal plans.

We now present our objective function. We distinguish between how the agent perceives the present which is certain and the future which is uncertain. Let $F_t : X_t \times Y_t \times Z_t \to R$ denote the payoff in period $t$ to the period $t$ individual resulting from the state $(x_t, z_t)$ and the action $y_t$. For $t < T$, let $\Phi_t : P_{t+1} \to R$ denote the value to the period $t$ individual of a given probability sequence over future states and acts. For period $T$ the agent simply maximizes the value of $F_T$. For $t < T$, the agent maximizes the expected value of the sum of the two terms:

$$F_t(x_t, y_t, z_t) + \Phi_t(p_{t+1}). \quad (3.1)$$

We place the following restrictions on the $F_t$ and the $\Phi_t$.

**Assumption 3.** For all $t$, $F_t$ and $\Phi_t$ are continuous in all of their arguments.\(^6\)

The dependence of $F$ and $\Phi$ on $t$ allows tastes to vary over time. If we let $F_{r,t} : X_r \times Y_r \times Z_r \to R$ denote the period $t$ view of the payoff in period $r$ resulting from the state $(x_r, z_r)$ and the action $y_r$ and take

$$\Phi_t(p_{t+1}) = E_{p_{t+1}} \sum_{r=t+1}^T F_{r,t}(x_r, y_r, z_r), \quad (3.2)$$

\(^5\)In an earlier version of this paper we derived Assumption 2 by taking a standard decision problem under uncertainty and introducing a shock to $x_{t+1}$ that is imperfectly correlated with $z_{t+1}$.

\(^6\)The spaces are compact, so continuity implies boundedness.
then maximizing the expected value of (3.1) reduces to maximizing the expected present value of payoffs from the perspective of period $t$. The objective (3.1), however, allows the agent a completely general response to uncertainty.

Given the time inconsistency of preferences, the determination of the appropriate strategy space is a very subtle matter, to which we now turn.

4. Plans

4.1. Definition: continuation plans

In the standard dynamic programming model, the agent is content to decide only today’s action, secure in the knowledge that when the time comes future actions will be chosen optimally. When tastes change, the recursive solution is a bit more complex. We have seen in the examples that choosing the best plan that will actually be followed involves the agent choosing today how to break indifference tomorrow. The recursive solution therefore calls for the agent to choose a plan for the entire future subject to the constraints imposed by the system and future choices. Given that the payoff function depends on the probability distribution over future states and actions, the agent may find it optimal to choose a plan which involves randomization in the face of future indifference.

The natural domain of choice which captures the agent’s concern over the present and the future as well as the agent’s concern over the probability distribution over future states and choices is $Y_t \times P_{t+1}$. We therefore define a plan in period $t$ as an act today and the choice of a probability distribution over future states and actions, $\pi_t = (y_t, p_{t+1}) \in Y_t \times P_{t+1}$.

**Definition.** For all $t \leq T$, a period $t$ plan, $\pi_t$, is an element of $Y_t \times P_{t+1}$. A plan $\pi$ is a plan from the initial period.

Of course, not all plans are consistent with the data of the system or with future behavior. We address these issues in turn.

4.2. Feasible choices and plans

The set of feasible plans are those that are consistent both with the constraints on actions and with the law of motion of the state variables. We define the set of feasible period $t$ plans recursively. In period $T$, a plan $\pi_T = (y_T, p_{T+1})$ is feasible given the state $s_T$, if $y_T \in \Pi_T(s_T)$ (recall that $p_{T+1}$ is a place holder included for notational convenience). Let $\Pi_T(s_T)$ denote the set of period $T$ plans feasible from the state $s_T$.

Given $\Pi_{t+1}(s_{t+1})$, we define $\Pi_t(s_t)$ by induction. Consider any plan $\pi_t = (y_t, p_{t+1})$. It is clear that a necessary condition for $\pi_t$ to be feasible is $y_t \in \Pi_t(s_t)$. The restrictions that feasibility place on $p_{t+1}$ are more subtle. Recall that $p_{t+1}$ is a probability distribution over $S_{t+1} \times Y_{t+1} \times P_{t+2}$. The probabilities that $p_{t+1}$ assigns to Borel sets in $S_{t+1}$ must correspond to the law of motion $v_{y_t, z_t}$. Moreover, given any realization of $s_{t+1}$, the future choices must be feasible. This implies that $p_{t+1}$ must not assign positive probability to Borel sets in $Y_{t+1} \times P_{t+2}$ that lie outside of $\Pi_{t+1}(s_{t+1})$. This last statement need not apply
at every $s_{t+1}$, but almost everywhere according to the measure $v_{y_t,z_t}$. This suggests that
$p_{t+1}$ is feasible if it can be represented by $\mu_{s_{t+1}} v_{y_t,z_t}$, where $\mu_{s_{t+1}}$ is a regular version of
the conditional probability of $(y_{t+1}, p_{t+2})$ given $s_{t+1}$, and where $\Pi_{t+1}(s_{t+1})$ is a support of
$\mu_{s_{t+1}} v_{y_t,z_t}$-almost everywhere. It will be convenient to define $\Lambda^f_t(y_T, z_T) = P_{T+1}$ and
for $t < T$ (where the superscript $f$ connotes feasible),
$$\Lambda^f_t(y_t, z_t) = \{p_{t+1} \in P_{t+1} \text{ such that}$$
$$p_{t+1} = \mu_{s_{t+1}} v_{y_t,z_t},$$
$$\mu_{s_{t+1}} \text{ is a regular version of the conditional probability}$$
of $y_{t+1}$ and $p_{t+2}$ given $s_{t+1}$, and
$$\Pi_{t+1}(s_{t+1}) \text{ is a support of } \mu_{s_{t+1}} v_{y_t,z_t} \text{-almost everywhere}\}$$
This leads to the following definition of a feasible choice:

**Definition.** A period $t$ plan $\pi_t = (y_t, p_{t+1}) \in Y_t \times P_{t+1}$ is feasible given $s_t$, if $y_t \in \Gamma_t(s_t)$
and $p_{t+1} \in \Lambda^f_t(y_t, z_t)$.

Lemma 1 contained in the appendix asserts that with Assumptions 1 and 2, the set of feasible plans is non-empty.

### 4.3. Recursively optimal plans and existence

The key qualitative impact of time inconsistency is that we can no longer treat all feasible plans as available to the decision maker. The concept of recursive optimality involves selecting plans that are not dominated in future states. We define the set of optimal plans recursively.

The definition of optimality is trivial in the final period. Given the state $s_T$, the agent chooses $\pi_T = (y_T, p_{T+1}) \in \Pi_T(s_T)$ to maximize $F_T(x_T, y_T, z_T)$. Let
$$V_T(s_T) \equiv \sup_{\pi_T = (y_T, p_{T+1}) \in \Pi_T(s_T)} F_T(x_T, y_T, z_T).$$

Let $G_T(s_T)$ denote the set of optimal plans in period $T$ given $s_T$:
$$G_T(s_T) \equiv \{\pi_T = (y_T, p_{T+1}) \in \Pi_T(s_T) | F_T(x_T, y_T, z_T) = V_T(s_T)\}.$$ (4.1)

In periods prior to $T$, the optimal plan will have to be consistent with optimality in the future. Let $G_{t+1}(s_{t+1})$ denote the set of recursively optimal period $t+1$ plans given $s_{t+1}$. In period $t$, the agent chooses $\pi_t = (y_t, p_{t+1})$ given the state $s_t$. This choice is constrained by feasibility, $\pi_t \in \Pi_t(s_t)$. It is also constrained by future behavior. Let $\Lambda^f_t(y_t, z_t) \subseteq \Lambda^f_t(y_t, z_t)$ denote the set of $p_{t+1} \in P_{t+1}$ that are feasible and consistent with future optimality, so that $p_{t+1}$ is supported by $G_{t+1}(s_{t+1})$ for $v_{y_t,z_t}$-almost every $s_{t+1}$:
$$\Lambda^f_t(y_t, z_t) = \{p_{t+1} \in \Lambda^f_t(y_t, z_t) \text{ such that}$$
$$p_{t+1} = \mu_{s_{t+1}} v_{y_t,z_t},$$
\[ \mu_{s_{t+1}} \text{ is a regular version of the conditional probability} \]

of \( y_{t+1} \) and \( p_{t+2} \) given \( s_{t+1} \), and

\[ G_{t+1}(s_{t+1}) \text{ is a support of } \mu_{s_{t+1}} \text{ } v_{y_{t}, z_{t}} \text{-almost everywhere}. \]  

(4.2)

Let \( \hat{\Pi}_t(s_t) \) denote the set of consistent period-\( t \) plans given \( s_t \). These are the choices that are feasible and consistent with future optimality:

\[ \hat{\Pi}_t(s_t) = \{(y_t, p_{t+1}) | y_t \in \Gamma_t(s_t) \text{ and } p_{t+1} \in A^c_t(y_t, z_t)\}. \]

The recursively optimal choices in period \( t \) are those that are consistent and maximize (3.1). Again let

\[ V_t(s_t) = \sup_{\pi_t \in \hat{\Pi}_t(s_t)} \left[ F_t(x_t, y_t, z_t) + \Phi_t(p_{t+1}) \right], \]

so that the set of recursively optimal period \( t \) plans given \( s_t \) is

\[ G_t(s_t) = \left\{ \pi_t = (y_t, p_{t+1}) \in \hat{\Pi}_t(s_t) \mid F_t(x_t, y_t, z_t) + \Phi_t(p_{t+1}) = V_t(s_t) \right\}. \]

Overall, a recursively optimal plan is an element of \( G_0(s_0) \).

The following proposition establishes the existence of such a plan. The proof is in the appendix.

**Proposition 4.1.** Given Assumptions 1–3, there exists a recursively optimal plan.

5. Discussion

5.1. The objective function

One significant feature of the theorem is that we are able to allow for a more general objective function than is typically considered in dynamic programming. Rather than taking the period \( t \) payoff to be the expected value of the payoff to all pure current and future states of nature, \( E_t[\sum_{r=1}^{T} F_{r,t}(x_r, y_r, z_r)] \), we allow the payoff to beliefs about future states to be quite general, \( \Phi_t(p_{t+1}) \). In essence, linearity of the payoff in beliefs about the future is discarded.

Kreps and Porteus [17] also allow for non-linearities in probabilities in their model of the preference over the temporal resolution of uncertainty. They assume, however, that preferences are time consistent. One way to view the objective (3.1) is as a generalization of Kreps–Porteus in which their assumption of time inconsistency has been relaxed and an additive separability assumption has been added.

By allowing for both non-linearities and time inconsistency, we open up a whole range of phenomena that are excluded in the standard set-up. The possibility of a stock market collapse may foster anxiety. Watching a sporting event may be more enjoyable while the outcome is in doubt. Anticipation of a pleasurable event may be more enjoyable if the agent is certain that the event will occur.
Caplin and Leahy [7] model these feelings with a payoff function of the form

\[ V_t(s_t) = F_t(x_t, y_t, z_t) + A_t(p_{t+1}) + E_t V_{t+1}(s_{t+1}). \]

Here \( A_t \) captures the influence of anticipation. The properties of this function depend on the type of anticipation being modeled, whether fear, suspense or anxiety. Note that since \( A_t \) enters the period \( t \) objective but not the period \( t+1 \) objective, so that \( A_t \) is sunk at the beginning of period \( t+1 \). This is the source of time inconsistency in their model.

More generally, our approach opens up the modeling of any dynamic decision problem with belief-based utility. Loomes and Sugden’s [21] model the disappointment and elation that results when outcomes differ from expectations. Geanakoplos et al. [14] and Caplin and Leahy [8] model games in which agents have belief-based utility functions that incorporate surprise and anxiety.

5.2. The nature of the optimal strategies

Our formulation of the model implies that what we ultimately prove is the existence of an optimal strategy in the “current action and future probability distribution” sense. Note that one can then step back to the classical concept of a full state-contingent plan and find an optimal strategy in the standard sense, although this plan may involve randomization.

5.3. Continuity and noise

The key issues in the proof concern inheritance of continuity properties. The main technical problem is the potential discontinuity of the value of an optimal policy tomorrow from today’s perspective. This potential discontinuity arises because of the consistency requirement. In choosing an optimal policy the agent must choose from those policies that will be optimal tomorrow. This latter set, however, is not a continuous correspondence of the future state, and so the theorem of the maximum, in its standard form, fails.

The continuity problems are alleviated by Assumption 2, which introduces noise into the relationship between current choices and future states. The noise “smooths” the ensuing state so that continuation values are continuous. It is qualitatively similar to the approach of Bernheim and Ray [6], who add noise to the return to saving for a similar reason in their proof of the existence of a Markov perfect equilibrium. Our setting is much more general in terms of both preferences and technology than the consumption-saving problem of Bernheim and Ray. Moreover, the recursive approach differs fundamentally from the Markov perfect approach in its treatment of ties. In the recursive approach, ties are broken by the strict preference of earlier selves, whereas choice in the Markov perfect approach must be independent of history.

As a technical note on the proof, the assumption of additive separability plays a significant role. If \( y_t \) and \( s_t \) had entered \( \Phi_t \) directly without the intermediation of the transition function \( \varphi \), then our source of noise would not have been sufficient to get the integrands in Lemma 3 to converge. One way to introduce time inseparability would be to have a gap between the intended choice and the realized choice. We could assume that the agent who chooses \( y_t \) does so with error, so that the variable that enters the payoff functions is \( y_t + \varepsilon_t \), where \( \varepsilon_t \) is
an error term.\textsuperscript{7} The proof of existence would be similar to that presented above. The case in which $\Phi_t$ depends on $s_t$ could be handled similarly. One either assumes that the state enters the payoffs with an error, or that the agent perceives the state imperfectly, so that the decision depends on the expectation of $s_t$. In either case the proof would proceed in a similar manner.

5.4. The noiseless limit

One potential use of the existence result is to apply noise to models in which a solution does not exist. In this subsection we present conjectures regarding the noiseless limit of the model.

We have in mind a sequence of models indexed by $n \in \{1, \ldots, \infty\}$. Each model is identical except that for $\{\lambda_{t,n}\}_{t=0}^{T}$ converges to a distribution with mass 1 at $x_{t+1} = y_t$, so that, in the limit, this period's choice becomes the next period's endogenous state. We want to understand the relationship between the sequence of “noisy” models with uncertain $x_{t+1}$ and the limiting model in which $x_{t+1}$ is deterministic.

To fix ideas, we discuss the limiting model in greater detail. Since $\lambda_t$ is degenerate in the limit and since $q_t$ is exogenous, choosing $y_t$ and $p_{t+1}$ is equivalent to choosing $y_t \in \Gamma(s_t)$ and a $\mu_{s_{t+1}} \in P(G_{t+1}(s_{t+1}))$ for each $z_{t+1}$, where $P(G_{t+1}(s_{t+1}))$ is the space of probability distributions over $G_{t+1}(s_{t+1})$. We can write value of an optimal policy in the limiting model as

$$\tilde{V}_t(s_t) \equiv \sup_{y_t \in \Gamma(s_t)} \mu_{t+1} \in P(G_{t+1}(s_{t+1})) \right \} F_t(x_t, y_t, z_t) + \Phi_t(p_{t+1}(y_t, q_t, \mu_{t+1})).$$

To keep things simple, we will consider the more standard case of time changing tastes in which $\Phi_t$ satisfies Eq. (3.2). For what follows it will be convenient to write this as a two-stage decision problem

$$\tilde{V}_t(s_t) \equiv \sup_{y_t \in \Gamma(s_t)} F_t(x_t, y_t, z_t) + E_t W_t(y_t, z_{t+1}),$$

where

$$W_t(y_t, z_{t+1}) = \sup_{\mu_{t+1} \in P(G_{t+1}(s_{t+1}))} E_{t+1} \sum_{r=t+1}^{T} F_{r,t}(x_r, y_r, z_r).$$

One question that naturally arises in this setting is when the strategies that are recursively optimal in the sequence of models converge to the strategy that is recursively optimal in the limiting model in the case in which the latter strategy exists. There are two cases in which such convergence may be a problem. The first arises when there are multiple policies that maximize $\tilde{V}_t(s_t)$ in the limiting model. In this case, different sequences of noisy models that converge to the same limiting model may favor different limiting policies. It is therefore proper to compare the optimal policy in the limiting model to the set of policies that are

\textsuperscript{7} Ameriks et al. [1] model a consumer who makes errors in consumption due to absentmindedness.
limits of any sequence of noisy models converging to that limiting model. The second case arises when there is a point $\bar{y}$ such that $W_t(\bar{y}, z_{t+1}) > \lim W_t(y_n, z_{t+1})$ for all $y_n \to \bar{y}$. This may happen if $G_{t+1}(s_{t+1})$ is not lower hemi-continuous from all directions at $x_{t+1} = \bar{y}$. In this case it is possible that $\bar{y}$ is the optimal choice in the limiting model. Since $x_{t+1} = \bar{y}$ is measure zero in all of the “noisy” models close to the limiting model, however, the value of a choice near $\bar{y}$ is discretely lower in these models. The optimal choice therefore may be very different. A condition that should be sufficient to eliminate this case is:

**Assumption 4.** For all $\bar{y}$, $W_t(\bar{y}, z_{t+1}) \leq \lim_{\epsilon \to 0} \max_{y_{\epsilon} \in N_\epsilon(\bar{y}) \setminus \bar{y}} W_t(y_{\epsilon}, z_{t+1})$.

Here $N_\epsilon(\bar{y}) \setminus \bar{y}$ is the complement of $\bar{y}$ in a neighborhood $\bar{y}$ of radius $\epsilon$. Assumption 4 is weaker than requiring that $W_t$ be lower semi-continuous, and neither stronger nor weaker than upper hemi-continuity.

A second question involves the characteristics of the limit of the sequence of models in the case in which there does not exist an optimal policy in the limiting model. Assuming that the condition in Assumption 4 holds, then the value functions will converge to $\tilde{V}_t(s_t)$ and the policies will converge to the policy that would have attained the supremum had the value function been lower semi-continuous at the limiting policy. This solution should coincide with the limit of $\epsilon$-optimal solutions to the limiting model. In the example of Section 2.3, this would correspond to the value that the period 1 agent would receive if this agent were allowed to decide indifference in period 2. It is not clear that the last property generalizes.

### 6. The strategic approach revisited

Viewed in strategic terms, it is clear that the recursively optimal strategy, when it exists, is always a subgame perfect equilibrium to the intrapersonal game. Perfection follows from the requirement that the current self believes that all future selves will in all circumstances make choices that are optimal from their perspectives. The recursive approach can thus be viewed as providing a refinement of subgame perfection.

Our main criticism of the game theoretic approach to modeling time inconsistency was its inattention to the possibilities for communication and coordination that are available to an individual. The game theoretic literature is not, however, silent on these issues. The related notions of coalition-proofness and renegotiation-proofness are concerned with how groups of agents may at different times coordinate on equilibria that improve their collective payoffs. It is therefore interesting to compare our recursive solution with the refinements.

Coalition-proofness extends the concept of Nash equilibrium from unilateral deviations by individuals to coordinated deviations by groups of individuals. Unlike the related concept of strong Nash equilibrium, coalition-proofness does not allow arbitrary deviations, but only those that are themselves stable in the face of further deviations. We show by example that coalition-proofness and recursive optimality are entirely different concepts. Neither nests the other.

---

8 See [5] for a discussion of coalition-proof Nash equilibria. A related topic is renegotiation-proofness (see [4]) which is commonly applied in repeated game settings.
Consider Example 1 of Section 2. Both \{A, a, c\} and \{B, b, c\} are perfectly coalition-proof. This is because each is Pareto optimal and the set of Pareto optimal subgame perfect equilibria and the set of coalition-proof equilibria are identical in two player games. Only \{A, a, c\}, however, is recursively optimal. Hence there exist perfectly coalition-proof equilibria that are not recursively optimal.

The reason that the two concepts differ here is that recursion allows for the breaking of indifference in period 2 to the advantage of player one, whereas coalition-proofness requires that all members of the deviating coalition be strictly better off. In this sense coalition-proofness is more stable against certain types of deviations.

To illustrate the converse, we extend Example 1 to include a period zero player who chooses from the set \{u, d\} as in Fig. 2. If the zero player chooses u, then player one and player two play the subgame described in Example 1. If the zero player chooses d, then players one and two have no choice. The payoffs are listed at the terminal nodes with the payoff to the zero player listed first, followed by the payoffs to the period one and period two players, respectively. The payoffs are same as in Example 1, with the addition that the zero player receives 10 if \{u, B, c\} is played, 5 if d is played, and zero otherwise. Players one and two receive zero if d is played. Now it is easy to show that if players one and two play \{A, a, c\} in the subgame then player zero chooses d, whereas if these agents play \{B, b, c\} then player zero chooses u. Everyone is better off, however, in the case \{u, B, b, c\} and this is the perfectly coalition-proof Nash equilibrium. The recursive solution, however, is \{d, A, a, c\}. This example shows that the recursive optimum may not be coalition-proof.
The reason that the two concepts differ here is a corollary to the first example. The robustness of coalition-proofness to certain types of deviations, namely the breaking of ties, allows agents to support other types of deviations. In this sense coalition-proofness is less stable.

7. Concluding remarks

We have provided new arguments and results that favor the recursive approach to time inconsistency introduced by Strotz rather than the standard strategic approach.

Acknowledgements

We would like to thank Eduardo Engel, Chris House, John Leahy Sr. and an anonymous referee for helpful comments and discussions. Caplin thanks the C.V. Starr Center at NYU and Leahy thanks the National Science Foundation and the Sloan Foundation for financial support.

Appendix

Lemma 1. For all \( s_t \in S_t \), \( \Pi_t(s_t) \) is non-empty.

Proof. The proof is by induction. In period \( T \), consider an arbitrary \( s_T \in S_T \). \( \Pi_T(s_T) \) denotes the set of choices \( \pi_T = (y_T, p_{T+1}) \) such that \( y_T \) satisfies the constraint \( \Gamma_T(s_T) \) and \( p_{T+1} \in P_{T+1} \). That \( \Pi_T(s_T) \) is non-empty follows from Assumption 1 which states that \( \Gamma_T(s_T) \) is non-empty. Moreover, by Assumption 1, \( \Gamma_T(s_T) \) is upper semi-continuous. Hence there exists a measurable selection \( \gamma_T(s_T) \) from \( \Gamma_T(s_T) \). It follows that \( (\gamma_T(s_T), p_{T+1}) \) is a measurable selection from \( \Pi_T(s_T) \).

Now suppose that for all \( s_{t+1} \in S_{t+1} \), \( \Pi_{t+1}(s_{t+1}) \) is non-empty and that there exists a measurable selection \( \pi_{t+1}(s_{t+1}) \) from \( \Pi_{t+1}(s_{t+1}) \). Consider \( s_t \in S_t \). By Assumption 1, \( \Gamma_t(s_t) \) is non-empty. Select \( y \in \Gamma_t(s_t) \). Given Assumption 2, \( s_t \) and \( y \) determine a probability distribution over \( s_{t+1}, v_{s_t, y} \). Now by assumption there exists a measurable selection \( \pi_{t+1}(s_{t+1}) \) from \( \Pi_{t+1}(s_{t+1}) \). Given any Borel set \( A \in Y_{t+1} \times P_{t+2} \), define the measure \( \mu_{x_{t+1}}(A) = 1_{\pi_{x_{t+1}}(s_{t+1}) \in A} \) where \( 1_B \) is the indicator of the event \( B \). \( \mu_{x_{t+1}} \) is a regular version of the conditional probability of \( y_{t+1} \) and \( p_{t+2} \) given \( s_{t+1} \), and by construction \( \Pi_{t+1}(s_{t+1}) \) is a support of \( \mu_{x_{t+1}} v_{s_t, y} \)-almost everywhere. Let \( p = \mu_{x_{t+1}} v_{s_t, y} \times (y, p) \) be feasible by construction. It follows that \( \Pi_t(s_t) \) is non-empty. By Assumption 1, \( \Gamma_t(s_t) \) is upper semi-continuous. Hence there exists a measurable selection \( \gamma_t(s_t) \) from \( \Gamma_t(s_t) \). Moreover, \( v_{s_t, y} \) is a measurable function of \( y_t \) and \( z_t \). It follows that \( (\gamma_t(s_t), \mu_{x_{t+1}} v_{\gamma_t(s_t), z_t}) \) is a measurable selection from \( \Pi_t(s_t) \). This completes the induction step and the proof of the lemma. □

Proposition A.1. Given Assumptions 1–3, there exists a recursively optimal plan.
Proof. Supporting lemmas follow. The proof is by induction. Given $z_T \in Z_T$, Assumptions 1 and 3 allow us to apply the standard theorem of the maximum [26, p. 62] in period $T$ to show that the optimal choice correspondence $G_T(s_T)$ in Eq. (4.1) is non-empty, compact valued, and upper hemi-continuous in $x_t$ (hence $G_T(\cdot, z_T)$ has a closed graph).

We now turn to the induction step. We assume that for each $z_{t+1} \in Z_{t+1}$, $G_{t+1}(s_{t+1})$, which gives $y_{t+1}$ and $p_{t+2}$ as a function of $s_{t+1}$, is a non-empty, compact valued, and upper hemi-continuous correspondence in $x_{t+1}$. We show that this implies that for each $z_t \in Z_t$, $G_t(s_t)$ is non-empty, compact valued, and upper hemi-continuous in $x_t$.

We divide the optimization problem (4.3) into two steps.

$$V_t(s_t) = \sup_{y \in \Gamma_t(s_t)} \left\{ F_t(x_t, y, z_t) + \sup_{p \in \Lambda^c_t(y, z_t)} \Phi_t(p) \right\}. $$

First, we consider the optimal choice of $p_{t+1}$ given $y_t$ and $s_t$. We then consider the optimal choice of $y_t$.

Consider the problem of choosing an optimal $p_{t+1}$ given $y_t$ and $z_t$:

$$W(y_t, z_t) = \sup_{p \in \Lambda^c_t(y, z_t)} \Phi_t(p). \quad(A.1)$$

The choice set is $\Lambda^c_t(y_t, z_t)$.

By Lemma 3, $\Lambda^c_t(y_t, z_t)$ is non-empty, compact valued and continuous. By Assumption 3, $\Phi_t(p)$ is continuous. Hence by Lemma 4 (a theorem of the maximum for non-Euclidian spaces), the optimal choice correspondence is non-empty, compact valued and upper hemi-continuous, and the maximized value $W(y_t, z_t)$ is continuous. Let $L(y_t, z_t)$ denote the optimal choice correspondence for problem (A.1).

Consider now the problem of choosing $y_t$ given $z_t$ and the optimal contingent choice of $p_{t+1}$:

$$V_t(s_t) = \max_{y \in \Gamma_t(s_t)} F_t(s_t, y) + W(y, z_t). \quad(A.2)$$

$\Gamma_t$ is non-empty, compact valued and continuous by Assumption 1. $F_t$ is continuous by Assumption 3. $W$ is continuous by the argument of the last paragraph. Therefore, by the theorem of the maximum, the optimal choice correspondence is upper hemi-continuous and the maximized value $V_t(s_t)$ is continuous. Let $M(s_t)$ denote the optimal choice correspondence for problem (A.2).

We construct $G_t(s_t)$ from $L(y_t, z_t)$ and $M(s_t)$ as follows:

$$G_t(s_t) = \{(y, p) | y \in M(s_t) \text{ and } p \in L(y, z_t)\}.$$ 

That $G_t(s_t)$ is non-empty follows immediately from the fact that $L(y_t, z_t)$ and $M(s_t)$ are non-empty.

To show that $G_t(s_t)$ is compact valued, fix $s_t$, and consider $(y_n, p_n) \in G_t(s_t)$ with $(y_n, p_n)$ converging to $(y, p)$. Since $(y_n, p_n) \in G_t(s_t)$, we have $F_t(s_t, y_n) + \Phi_t(p_n) = c$. Since $F_t + \Phi_t$ is continuous, we have $F_t(s_t, y) + \Phi_t(p) = c$ as well. It follows that $(y, p) \in G_t(s_t)$. Since $G_t(s_t)$ contains all of its limits, it is closed. Since it is also a subset of $Y_t \times P_{t+1}$ which is compact, it is compact.
To show that $G_t(s_t)$ is upper hemi-continuous, fix $s \in S_t$. Consider $s_n \in S_t$ with $s_n \to s$, and $(y_n, p_n) \in G_t(s_n)$. Since $(y_n, p_n) \in Y_t \times P_{t+1}$, which is compact, there exists a convergent subsequence $(y_m, p_m)$ converging to $(y, p)$. Since $M(s_t)$ is upper hemi-continuous $y \in M(s_t)$, and since $L(y_t, z_t)$ is upper hemi-continuous $p \in L(y_t, z_t)$. It follows from the definition of $G_t(s)$ that $(y, p) \in G_t(s)$. This proves that $G_t(s)$ is upper hemi-continuous. This completes the induction step and the proof.\footnote{Some authors define a plan as a Borel measurable function of the current state. Given that $G_t(s)$ is upper hemi-continuous, we can apply the measurable selection theorem of Hinderer [16] to show that such a function exists.}

**Lemma 2.** Let $p$ be a distribution on $S_t \times Y_t \times P_{t+1}$. Then given Assumption 1 there exists a distribution $v$ on $S_t$ and a transition probability kernel $K$ from $S_t$ into $Y_t \times P_{t+1}$ such that

$$p(ds_t, dy_t, dp_{t+1}) = K_{s_t}(dy_t, dp_{t+1})v(ds_t).$$

**Proof.** According to Assumption 1, $S_t$ and $Y_t$ are compact subsets of finite-dimensional Euclidean spaces for all $t \in [0, T]$. $S_t$ and $Y_t$ are therefore compact separable metric spaces. It follows that $P_{t+1}$ is a compact separable metric space [17] and that $Y_t \times P_{t+1}$ is a compact separable metric space. By Dellacherie and Meyer [10, III-16-200], $Y_t \times P_{t+1}$ is a measurable Lusin space. The lemma follows from Çinlar [9, II.7.14].

**Lemma 3.** Let $G_{t+1}$ be a non-empty, compact valued correspondence that is upper hemi-continuous in $x_{t+1}$ and measurable in $z_{t+1}$. Then the correspondence $\Lambda^c_{t}(y_t, z_t)$ is continuous in $y_t$ and measurable in $z_t$.

**Proof.** We first show that $\Lambda^c_t(y_t, z_t)$ is non-empty. Consider a measurable selection from $G_{t+1}(x_{t+1}, z_{t+1})$, $g(x_{t+1}, z_{t+1})$, and for any measurable set $A \subseteq Y_{t+1} \times P_{t+2}$ define $\mu(A|s_{t+1}) = 1_{g(x_{t+1}, z_{t+1}) \in A}$. Set $p = v(y_t, z_t).$

We now show that $\Lambda^c_t(y_t, z_t)$ is compact valued. Fix $(y_t, z_t)$. Consider a sequence $p_n \in \Lambda^c_t(y_t, z_t)$ with $p_n$ converging to $p \in P_{t+1}$. By Lemma 2, there exist $\mu_n(dy_{t+1}, dp_{t+2}|s_{t+1})$ such that $p_n = \mu_n y_{t+1}, z_{t+1}$ and $\mu(dy_{t+1}, dp_{t+2}|s_{t+1})$ such that $p = \mu y_{t+1}, z_{t+1}$. We now show that $G_{t+1}(s_{t+1})$ is a support of $\mu y_{t+1}, z_{t+1}$-almost everywhere. Suppose not. Then there exists a set $A \subseteq S_{t+1}$ such that $v_{y_t, z_t}(A) > 0$ and for $s_{t+1} \in A$, $\mu(G^c_{t+1}(s_{t+1})) > 0$. Since the distribution of $z_{t+1}$ is simple, we may consider a single $z$-section of $A$: there exists a $\bar{z} \in Z_{t+1}$ and $A' = \{(y, \bar{z}) \in A\}$ such that $v_{y_t, \bar{z}}(A') > 0$ and for $s_{t+1} \in A'$, $\mu(G^c_{t+1}(s_{t+1})) > 0$. Let $B = \{(x_{t+1}, \bar{z}, y_{t+1}, p_{t+1})| (y_{t+1}, z_{t+1}) \in G_{t+1}(x_{t+1}, \bar{z})\}$ denote the graph of $G_{t+1}$ given $\bar{z}$. Since $G_{t+1}$ is compact valued and upper hemi-continuous, $B$ is closed. Now $p(B^c) > 0$. Since $p$ is regular [13, Theorem 7.8], there exists a compact set $D \subseteq B^c$ such that $p(D) > 0$. Since both $D$ and $B$ are closed and since $D \cap B = \emptyset$, there exists by Uryshon’s lemma a continuous function $f$ which takes the value of one on $D$ and zero on $B$. Note $\int f dp > 0$ and $\int f dp_0 = 0$ which contradicts weak convergence. This contradiction establishes that $p \in \Lambda^c_t(y_t, z_t)$. Since $\Lambda^c_t$ contains all of its limits it is closed. Since $\Lambda^c_t(y_t, z_t) \subseteq P_{t+1}$ and $P_{t+1}$ is compact, $\Lambda^c_t(y_t, z_t)$ is compact.

We now show that given $z \in Z_t$, $\Lambda^c_t(y, z)$ is upper hemi-continuous in $y$. Fix $z \in Z_t$ and consider $y_n$ converging to $y$ with $p_n \in \Lambda^c_t(y_n, z)$. Since $p_n \in \Lambda^c_t(y_n, z) \subseteq P_{t+1}$ and $P_{t+1}$
We show that $p_n$ converging to $p \in P_{t+1}$. The demonstration that $p \in \Lambda^c_t(y_t, z_t)$ follows the argument in the last paragraph.

We now show that given $z \in Z_t$, $\Lambda^c_t(y, z)$ is lower hemi-continuous in $y$. Consider $p \in \Lambda^c_t(y, z)$ and $y_n$ converging to $y$. Since $p \in \Lambda^c_t(y, z)$,

$$p(dx_{t+1}, dy_{t+1}, dp_{t+2}) = \mu_{st+1}(dy_{t+1}, dp_{t+2})v_{y,z}(dx_{t+1}, dz_{t+1})$$

for some $\mu_{st+1}(dy_{t+1}, dp_{t+2})$ and $G_{t+1}(s_{t+1})$ is a support of $\mu_{st+1}(dy_{t+1}, dp_{t+2})$ $v_{y,z}$-almost everywhere. We construct a sequence $p_n \in \Lambda^c_t(y_t, z_t)$ converging to $p$. First, amend $\mu_{st+1}$ as follows. Let $A$ denote the union of open boxes in $S_{t+1}$ which have strictly positive measure under the Borel $\sigma$-algebra on $S_{t+1}$ and which are measure zero according to $v_{y,z}$.

By Hinderer [16], there exists a measurable selection $\pi_{t+1}$ from $G_{t+1}(s_{t+1})$. Define $\hat{\mu}_{st+1}$ as follows:

$$\hat{\mu}_{st+1}(dy_{t+1}, dp_{t+2}) = \begin{cases} \mu_{st+1}(dy_{t+1}, dp_{t+2}) & \text{if } s_{t+1} \notin A, \\ I[\pi_{t+1}(s_{t+1})] & \text{if } s_{t+1} \in A, \end{cases}$$

where $I[\{B\}]$ is the indicator of the event $B$. Since $\hat{\mu}_{st+1}(dy_{t+1}, dp_{t+2}) = \mu_{st+1}(dy_{t+1}, dp_{t+2})$ on $A^c$, and since $v_{y,z}(A^c) = 1$, $G_{t+1}(s_{t+1})$ is a support of $\hat{\mu}_{st+1}(dy_{t+1}, dp_{t+2})$ $v_{y,z}$-almost everywhere.

Let

$$p_n(dx_{t+1}, dy_{t+1}, dp_{t+2}) = \hat{\mu}_{st+1}(dy_{t+1}, dp_{t+2})v_{y,z}(dx_{t+1}, dz_{t+1}).$$

We show that $p_n \in \Lambda^c_t(y_t, z_t)$. Suppose not. Then there exists a set $B \in S_{t+1}$ with $v_{y,z}(B) > 0$ such that $\hat{\mu}_{st+1}(G_{t+1}(s_{t+1})) < 1$. Since $v_{y,z}$ is absolutely continuous with respect to the Lebesgue measure on $S_{t+1}$, the Lebesgue measure of $B$ is strictly positive. Since $S_{t+1}$ is a Euclidean space $B$ is the countable union of boxes in $S_{t+1}$. Hence $B$ contains an open set whose Lebesgue measure is strictly positive. But this set must lie in $A$ and $\hat{\mu}_{st+1}(G_{t+1}(s_{t+1})) = 1$ on $A$. This contradiction establishes that $G_{t+1}(s_{t+1})$ is a support of $\hat{\mu}_{st+1}(dy_{t+1}, dp_{t+2})$ $v_{y,z}$-almost everywhere and that $p_n \in \Lambda^c_t(y_t, z_t)$.

We now show that $p_n$ converges to $p$ weakly. Let $f : S_{t+1} \times Y_{t+1} \times P_{t+2} \to R$ be bounded and continuous, and consider

$$\int f(s_{t+1}, y_{t+1}, p_{t+2})p_n(dx_{t+1}, dy_{t+1}, dp_{t+2}) = \int f(s_{t+1}, y_{t+1}, p_{t+2})\lambda(y_n, z_n, x_{t+1}, z_{t+1})\hat{\mu}_{st+1}(dy_{t+1}, dp_{t+2})dx_{t+1}dz_{t+1}.$$ 

By Assumption 2 the integrand is bounded. Hence $\int f dp_n$ converges to $\int f dp$ by the Lebesgue dominated convergence theorem, and $p_n$ converges to $p$ weakly. This establishes that $\Lambda^c_t(y_t, z_t)$ is lower hemi-continuous.

Since $\Lambda^c_t(y_t, z_t)$ is both lower and upper hemi-continuous, it is continuous. This completes the proof.

---

The issue here is that whereas $G_{t+1}(s_{t+1})$ is a support of $\mu_{st+1}(dy_{t+1}, dp_{t+2})$ $v_{y,z}$-almost everywhere, it may not be a support almost everywhere according to the Borel $\sigma$-algebra on $S_{t+1}$. 

10 The issue here is that whereas $G_{t+1}(s_{t+1})$ is a support of $\mu_{st+1}(dy_{t+1}, dp_{t+2})$ $v_{y,z}$-almost everywhere, it may not be a support almost everywhere according to the Borel $\sigma$-algebra on $S_{t+1}$. 


Lemma 4. Let $X$ and $Y$ be topological spaces and suppose that $Y$ is sequentially compact; let $f : X \times Y \to R$ be a continuous function; let $\Gamma : X \to Y$ be a non-empty, compact valued and continuous correspondence; for each $x \in X$, let $h(x) = \max_{y \in \Gamma(x)} f(x, y)$ and $G(x) = \{ y \in \Gamma(x) : f(x, y) = h(x) \}$. Then $h(x)$ is continuous and $G(x)$ is a non-empty, compact valued and upper hemi-continuous correspondence.

Proof. First we show that $G(x)$ is non-empty. Fix $x \in X$. The set $\Gamma(x)$ is non-empty and compact, and $f(x, \cdot)$ is continuous. Hence the maximum in $h(x)$ is attained and the set $G(x)$ is non-empty.

Now we show that $G(x)$ is compact valued. Fix $x$ and suppose $y_n \in G(x)$ such that $y_n \to y$. Since $\Gamma(x)$ is closed $y \in \Gamma(x)$. Since $y_n \in G(x)$, $f(x, y_n) = h(x)$. Since $f$ is continuous, $f(x, y) = h(x)$. It follows that $y \in G(x)$. Since $G(x)$ contains all of its limits it is closed. Since $\Gamma(x)$ is compact and $G(x)$ is a closed subset of $\Gamma(x)$, $G(x)$ is compact.

Next we show that $G(x)$ is upper hemi-continuous. Fix $x$, and let $\{x_n\}$ be any sequence converging to $x$. Choose $y_n \in G(x_n)$, all $n$. Since $Y$ is sequentially compact there exists a subsequence $\{y_m\}$ converging to $y$. Since $\Gamma$ is upper hemi-continuous $y \in \Gamma(x)$. Since $f$ is continuous, $\lim_{m \to \infty} f(x_m, y_m) = f(x, y)$. Let $z \in \Gamma(x)$. Since $\Gamma$ is lower hemi-continuous, there exists a sequence $z_m \to z$ with $z_m \in \Gamma(x_m)$ all $m$. Since $f(x_m, y_m) \geq f(x_m, z_m)$, all $m$, $f(x, y) \geq f(x, z)$. It follows that $y \in G(x)$ and $G(x)$ is upper hemi-continuous.

We now show that $h(x)$ is continuous. Fix $x \in X$ and $\{x_n\}$ be any sequence converging to $x$. Choose $y_n \in G(x_n)$, all $n$. Let $\tilde{h} = \lim sup h(x_n)$. There exists a subsequence $\{x_{n_k}\}$ such that $\tilde{h} = \lim f(x_{n_k}, y_{n_k})$. Since $Y$ is sequentially compact, there exists a subsequence of $\{y_{n_k}\}$, call it $\{y_j\}$, converging to $y$. Since $f$ is continuous $\lim f(x_j, y_j) = f(x, y) = \tilde{h}$. $G$ is upper hemi-continuous $y \in G(x)$. Hence $h(x) = \tilde{h}$. Repeat these arguments with $\lim inf h(x_n)$. It follows that $\lim inf h(x_n) = \lim sup h(x_n)$ and $h(x_n)$ is continuous. □

References