Recursive contracts in non-convex environments[*]

Chengfeng Shen[†]

Felix Kübler[‡]

Zhennan Zhou[§]

November 26, 2025

Abstract

In this paper we examine non-convex dynamic optimization problems with forward looking constraints. We prove that the recursive multiplier formulation in Marcet and Marimon (2019) gives the optimal value if one assumes that the planner has access to a public randomization device and forward looking constraints only have to hold in expectations. Whether one formulates the functional equation as a sup-inf problem or as an inf-sup problem is essential for the timing of the optimal lottery and for determining which constraints have to hold in expectations. We discuss for which economic problems the use of lotteries can be considered a reasonable assumption. We provide a general method to recover the optimal policy from a solution of the functional equation. As an application of our results, we consider the Ramsey problem of optimal government policy and give examples where lotteries are essential for the optimal solution.

JEL classification:

Keywords: Recursive contracts, lotteries, multiplier method, dual approach.

[†]Peking University. Email: shencf1999@pku.edu.cn

[‡]University of Zurich. Email: fkubler@gmail.com

[§]Westlake University. Email: zhouzhennan@westlake.edu.cn

1

1   Introduction

In many economic problems (for example, in contracting problems where agents are subject to intertemporal incentive constraints, or in models of optimal policy design where agents' reactions to government policies must be taken as constraints) the "natural state space" no longer suffices to describe optimal solutions and future promised utilities must be included in the state space (see, for example, Spear and Srivastava (1987)). This often makes the problem intractable for practical computations. Marcet and Marimon (2019) provide an alternative approach for a recursive formulation of a large class of dynamic models with forward-looking constraints which has been widely used in applied work (see, among many others, Cooley et al. (2004), Kehoe and Perri (2002), Attanasio and Rıos-Rull (2000), Aiyagari et al. (2002)). Despite these applications, there remain fundamental issues with respect to the applicability of the approach. In particular, Marcet and Marimon (2019) only consider the "convex case" where the objective function is concave and the constrained set convex. In many important applications, this assumption is violated. In fact, one of the two examples that Marcet and Marimon (2019) provide to illustrate the practical relevance of their method violates the assumptions – in the Ramsey model of optimal government policy, constraints consist of agents' first-order conditions and the feasible set is not convex. Since the theoretical analysis in Marcet and Marimon (2019) assumes the existence of a saddle point, one might fear that the method cannot be used for the important non-convex case, and the results in Pavoni et al. (2018a) and in Bloise and Siconolfi (2022) seem to confirm this fear.

In this paper, we consider lottery solutions to problems with forward-looking constraints and non-convexities and show that versions of the functional equation in Marcet and Marimon (2019) can be used to obtain the optimal value of the optimization problem if the decision maker has a public randomization device (lotteries) at his disposal. The crucial economic question then is in what scenarios it is reasonable to require the forward-looking constraints to hold only "in expectation". Forward looking constraints associated with time t typically impose restrictions on actions at periods $t, t+1, \ldots, t+N$, $1 \leq N \leq \infty$. While constraints in the current period typically need to hold for all outcomes of the lottery, it seems natural to assume that after the first period each forward looking constraint only needs to hold in expectations. In most applications, these are constraints on agents' future utility or marginal utility, and the assumption of expected utility then implies that under randomization the constraints should hold in expectations.

Non-convex optimization problems are challenging to solve even in the absence of forward-looking constraint. The numerical difficulties that arise have sparked substantial interest in exploring lottery solutions to models with non-convexities (see, e.g. Myerson (1982); Prescott and Townsend (1984); Stiglitz (1982) for classical contributions). In this approach, it is assumed that there is a public randomization device and decision makers can base their actions on the outcome of the randomization. This reformulation linearizes the constraints in the probability space and results in a convex problem.

Our paper makes four contributions. First, we prove that the dual problem considered in Pavoni et al. (2018a) gives the correct value if one allows for arbitrary lotteries over and all forward looking constraints need to hold only in expectations. Following Pavoni et al. (2018a), we show that the dual problem can be formulated as a recursive functional equation (FE), similar to the one postulated in Marcet and Marimon (2019) even if the problem is not convex. We will refer to this as the inf-sup FE.

Unfortunately, allowing for arbitrary lotteries often has no economic interpretation. In particular, we show that for the Ramsey problem of optimal government policy (Aiyagari et al. (2002)), the (full) lottery solution does not have a meaningful economic interpretation because it requires agents' first order condition to only hold on average. An economic interpretation can

be restored if agents can make optimal choices in the current period depending on the outcome of the lottery in the current period but if the, due to the assumption of expected utility, next period's marginal utilities have to be equal to today's in expectations. Our second contribution is to show that for this case, the dual formulation in Pavoni et al. (2018a) does not give the correct value of the problem but that if one modifies the functional equation from inf-sup to sup-inf, the correct (restricted) lottery solution is obtained. The resulting sup-inf FE can be analyzed with the same methods as the inf-sup FE.

As Cole and Kubler (2012) and Marimon and Werner (2021) point out, it can be difficult to recover the optimal (non-lottery) policy from the solution of the functional equation if the problem is not *strictly convex*. The introduction of lotteries transforms non-convex problems into convex ones, but obviously not into strictly convex problems. Our third contribution is to develop a computational method to recover the optimal policy. The basic idea is based on the insight in Marimon and Werner (2021) that establishes a relation between the subdifferential of the value function and possible utility promises. However, unlike Marimon and Werner (2021) (who focus on the simple case without uncertainty) we provide an algorithm that chooses a promise from each subdifferential that yields the highest value of the objective function.

As an application of our results, we consider a Ramsey problem of optimal government policy with incomplete markets (as in Aiyagari et al. (2002)). The problem is obviously not convex – however, that does not imply that for all specifications of parameters the optimal solution involves lotteries. The fourth contribution of our paper is to give an example to illustrate that lotteries might be crucial for the optimal solution in the Ramsey problem. In this example, the solution to the inf-sup functional equation is meaningless and the solution to the sup-inf equation gives the correct value of the lottery solution where the government can randomize over tax rates and government debt.

The remainder of the paper is organized as follows. After reviewing the existing literature, we provide a simple example to illustrate the main ideas in Section 2. Section 3 gives a general formulation of the problem and a recursive characterization of its dual. Section 4 links the value of the dual problem to lotteries and shows that the sup-inf version of the functional equation gives the optimal value of a related, often economically more relevant, lottery problem. Section 5 discusses how to recover optimal policies, Section 6 considers the Ramsey model from Aiyagari et al. (2002) as an application.

1.1 Related Literature

Following Spear and Srivastava (1987), the standard method to recursively solve dynamic contracting problems treated the promised utility of the agent(s) as a state variable, and to use this state variable to capture the implicit prior promises. Marcet and Marimon (2019) (the first working paper version was dated 1994) were the first to consider using Lagrange multipliers as part of the state. Pavoni et al. (2018a) show how the dual problem can be characterized by a functional equation even in the absence of convexity. They are also the first to prove the existence of Lagrange multipliers (in an appropriate space) for the problem.

Cole and Kubler (2012) and Marimon and Werner (2021) discuss how optimal policies can be recovered from a solution to the function equation. As we explain in detail below, our method to recover the optimal policy is based on their work.

Bloise and Siconolfi (2022) offer a seemingly alternative method to the problem based on the promised value approach. Although it is not clear whether their method can be made computationally practical, they show that in many scenarios it requires fewer assumptions about randomization than the dual approach by Pavoni et al. (2018a). As we show in Appendix D, in our setup their approach turns out to be equivalent to considering the sup-inf functional equation.

3

There is a large literature on the use of lotteries in static non-convex problems. Lotteries may arise in constrained optimal allocations whenever non-convexities are present, as shown in applications ranging from stochastic tax schedules to labor market welfare programs (e.g., Weiss (1976); Arnott and Stiglitz (1988); Pavoni and Violante (2007); Gauthier and Laroque (2014)). Shen et al. (2025) are the first to exploit the equivalence of the value of the dual (Lagrangian) problem and the lottery problem in economics and develop an efficient computational method to approximate optimal lotteries numerically.

2   A simple example

To illustrate the main results of the paper, we first present a very simple example. Suppose a principal maximizes

$$V^0 = \max_{(c_t, l_t)_{t=0}^\infty} \sum_{t=0}^\infty \beta^t (l_t - c_t), \ l_t \in [0,1], c_t \geq 0 \ \forall t$$

subject to a participation constraint of an agent who has non-convex preferences over consumption and labor,

$$\sum_{t=0}^\infty \beta^t ({c_t}^\sigma - {l_t}^\sigma) \geq 0, \ \sigma \in (0,1) \tag{2.1}$$

Following Marcet and Marimon (2019), one can imagine the following two recursive formulations of the problem which we will refer to as the sup-inf and the inf-sup functional equations[1].

$$V^1 = \max_{c,l} \min_\gamma \left[(l-c) + \gamma(c^\sigma - l^\sigma) + \beta W(\gamma) \right],$$

or

$$V^2 = \min_\gamma \max_{c,l} \left[(l-c) + \gamma(c^\sigma - l^\sigma) + \beta W(\gamma) \right],$$

where $W(\gamma)$ is a solution to

$$W(\gamma) = \max_{c,l} \left[(l-c) + \gamma(c^\sigma - l^\sigma) + \beta W(\gamma) \right].$$

In this section, we consider two examples for σ, β. In the first example, we have $V^0 = V^1 < V^2$, in the second example we have $V^0 < V^1 = V^2$. These examples seem to cast doubt on the usefulness of the recursive multiplier method in a non-convex problem. However, the main result of this paper shows that both V^1 and V^2 are the correct value of the maximization problem if the principal has access to a public randomization device (which we refer to as lotteries). The difference between V^1 and V^2 is due to a subtle difference in the timing between the realization of the lottery and the agent having to agree to participate in the contract. To make this clear, let $\mathcal{P}^\infty$ denote the space of probability measures on $([0,1] \times \mathbb{R}_+)^\infty$. One can imagine a contract that involves (independent) lotteries every period and the agent observes the first realization of that lottery before entering the contract. We refer to this case as "ex-post" lotteries. The value of the principal is given by[2]

$$V^{ep} = \max_{P \in \mathcal{P}^\infty} E^{(c_t, l_t)_{t=0}^\infty \sim P} \sum_{t=0}^\infty \beta^t (l_t - c_t),$$

---

[1]In this simple example, they are very simple functional equations in $W(.)$ – for the general case this naming will prove more accurate.

[2]The notation $c_0, l_0 \in \text{supp}(P)$ indicates that the pair (c_0, l_0) belongs to the support of P's marginal distribution on (c_0, l_0).

$$\text{s.t. } (c_0{}^\sigma - l_0{}^\sigma) + E^{(c_t,l_t)_{t=1}^\infty \sim P} \sum_{t=1}^\infty \beta^t (c_t{}^\sigma - l_t{}^\sigma) \geq 0 \text{ for all } c_0, l_0 \in \text{supp}(P).$$

Alternatively, one could assume that the agent has to agree to the contract before observing the first realization of the lottery. We refer to this as "ex-ante" lottery. The optimal value in this case is given by

$$V^{ea} = \max_{P \in \mathcal{P}^\infty} E^{(c_t,l_0)_{t=0}^\infty \sim P} \sum_{t=0}^\infty \beta^t (l_t - c_t),$$

$$\text{s.t. } E^{(c_t,l_t)_{t=0}^\infty \sim P} \sum_{t=1}^\infty \beta^t (c_t{}^\sigma - l_t{}^\sigma) \geq 0,$$

In this simple example, the difference between ex ante and ex post lotteries seems benign. However, we will argue below that in many applications ex ante lotteries do not have a sensible economic interpretation while ex post contracts do.

In Theorem 4.5 we will prove a general result that implies $V^2 = V^{ea}$ and in Theorem 4.6 we will prove a result that implies $V^1 = V^{ep}$. The two variations of the functional equations of Marcet and Marimon (2019) have no relation to the original problem (lottery-free), but give the correct value when the planner has access to a randomization device. The following example gives a simple intuition for this result.

2.1 Example 1

Suppose $\sigma = \frac{1}{2}$ and $\beta < \frac{1}{2}$. As we shall verify in the following, an optimal lottery-free solution is given by $l_0 = 1$, and for $t > 0$, $l_t = 0$, and to ensure that the participation constraint holds, $\frac{1}{1-\beta}\sqrt{c_t} = 1$, i.e. $c_t = (1-\beta)^2$ for all t. This gives the principle a value of $V^0 = \beta$. It is easy to see that the use of ex-ante lotteries can improve upon this solution. The optimal ex ante lottery will turn out to be $l_t \in \{0, 1\}$ with equal probability every period. With $c_t = 0.25$ for all t the (PC) is always satisfied ex ante. The value of the principal's problem is $0.25\frac{1}{1-\beta}$ which is larger than β whenever $\beta < \frac{1}{2}$.

The optimal value of this lottery solution can be obtained form the following inf sup FE.

$$V^2 = \min_\gamma \max_{c,l} \left[(l - c) + \gamma(\sqrt{c} - \sqrt{l}) + \beta W(\gamma) \right],$$

where $W(\gamma)$ solves

$$W(\gamma) = \max_{c,l} \left[(l - c) + \gamma(\sqrt{c} - \sqrt{l}) + \beta W(\gamma) \right].$$

It is easy to see that

$$W(\gamma) = \begin{cases} \frac{1}{1-\beta}\left(1 - \gamma + \frac{\gamma^2}{4}\right), & \gamma \leq 1 \\ \frac{1}{1-\beta}\frac{\gamma^2}{4} & \gamma > 1 \end{cases},$$

and that there are two optimal solutions, $(\gamma^1, c^1, l^1) = (1, \frac{1}{4}, 0)$, $(\gamma^2, c^2, l^2) = (1, \frac{1}{4}, 1)$, both giving a value of $V^2 = \frac{0.25}{1-\beta}$. As Shen et al. (2025) show, it is not a coincidence that this is the same as the value of the (ex-ante) lottery solution to the original problem: It is easy to see that

$$V^2 = \min_\gamma \max_{P \in \mathcal{P}} E^{(c,l) \sim P} \left[(l - c) + \gamma(\sqrt{c} - \sqrt{l}) + \beta W(\gamma) \right],$$

since the optimal lottery only puts weight on actions that give the same value. Since the problem is linear in probabilities, duality holds, i.e.

$$V^2 = \max_{P \in \mathcal{P}} \min_\gamma E^{(c,l) \sim P} \left[(l - c) + \gamma(\sqrt{c} - \sqrt{l}) + \beta W(\gamma) \right].$$

5

But for the (linearized) lottery problem, the results in Marcet and Marimon (2019) hold and we must have $V^2 = V^{ea}$.

The sup-inf FE turns out to be a bit more subtle. We have

$$V^1 = \max_{c,l} \min_\gamma \left[(l-c) + \gamma(\sqrt{c} - \sqrt{l}) + \beta W(\gamma) \right],$$

with $W(\gamma)$ as above. Then it follows that for $\gamma \geq 1$, we have $\gamma(c,l) = (\sqrt{l} - \sqrt{c})2\frac{1-\beta}{\beta}$ and $c = (1-\beta)^2, l = 1$, maximizes $(l-c) + \gamma(c,l)(\sqrt{c} - \sqrt{l}) + \beta W(\gamma(c,l))$ and $\gamma = 2(1-\beta)$ minimizes the expression for $c = (1-\beta)^2, l = 1$. Hence we obtain $V^1 = \beta$. As is shown in Theorem 4.6 this is the optimal ex post lottery solution which in this case coincides with a feasible lottery-free allocation that therefore must also be optimal. Hence we must have $V^0 = V^1$. This also verifies that the optimal lottery-free solution has a value of $V^0 = \beta$ since its value cannot exceed the one of the ex-post lottery solution. Therefore, in this example we have $V^2 > V^1 = V^0$.

2.2 Example 2

Now suppose that

$$0 < \beta < 1/2, \; 1 - \frac{\beta}{2} = \sigma^{\frac{\sigma}{1-\sigma}} \tag{2.2}$$

Note that in the previous case where $\sigma = 1/2$, $\sigma^{\sigma/(1-\sigma)} = 1/2$ and there is no $\beta < 1/2$ that satisfies the condition. On the other hand, when $\sigma \to 0, \sigma^{\frac{\sigma}{1-\sigma}} \to 1$ and we can find $\sigma \in (0, 1/2)$ and $\beta < 1/2$ to ensure that (2.2) is satisfied.

Similarly to the above, we obtain

$$W(\gamma) = \begin{cases} \frac{1}{1-\beta} \left[(\sigma^{\frac{\sigma}{1-\sigma}} - \sigma^{\frac{1}{1-\sigma}})\gamma^{\frac{1}{1-\sigma}} + (1-\gamma) \right], & \gamma \leq 1; \\ \frac{1}{1-\beta} (\sigma^{\frac{\sigma}{1-\sigma}} - \sigma^{\frac{1}{1-\sigma}})\gamma^{\frac{1}{1-\sigma}}, & \gamma > 1. \end{cases}$$

Regarding the problem for V, it is easy to verify that $(c,l) = (\sigma^{\frac{1}{1-\sigma}}, 1), \gamma = 1$ is a saddle point: Given $\gamma = 1$, we have that $(c,l) = (\sigma^{\frac{1}{1-\sigma}}, 1)$ is a maximizer; given $(c,l) = (\sigma^{\frac{1}{1-\sigma}}, 1)$, defining $F(\gamma) = (l-c) + \gamma(c^\sigma - l^\sigma) + \beta W(\gamma)$, we obtain

$$V = \frac{1}{1-\beta}(1 - \frac{\beta}{2} - \sigma^{\frac{1}{1-\sigma}}),$$

and one lottery solution is

$$l_0 = 1, p(l_t = 0) = p(l_t = 1) = 1/2 (t \geq 1), c_t \equiv \sigma^{\frac{1}{1-\sigma}}.$$

In this example, ex-ante and ex-post lotteries are the same because in the max-min problem for V we find a saddle point. On the other hand, it is easy to show that the deterministic (lottery-free) solution will not have the same value as the two lottery solutions, i.e. $V^0 < V^1 = V^2$. We have

$$\sum_{t=0}^\infty \beta^t(l_t - c_t) \quad \leq \sum_{t=0}^\infty \beta^t l_t^\sigma - \sum_{t=0}^\infty c_t \qquad \leq \sum_{t=0}^\infty \beta^t c_t^\sigma - \sum_{t=0}^\infty c_t$$

$$= \sum_{t=0}^\infty \beta^t(c_t^\sigma - c_t) \quad \leq \sum_{t=0}^\infty \beta^t(\sigma^{\frac{\sigma}{1-\sigma}} - \sigma^{\frac{1}{1-\sigma}}) \quad = \frac{1}{1-\beta}(1 - \frac{\beta}{2} - \sigma^{\frac{1}{1-\sigma}}).$$

For all inequalities to hold as equalities, we must have $l_t \in \{0, 1\}$ for all t and

$$\sum_{t=0}^\infty \beta^t l_t = \frac{1}{1-\beta}(1 - \frac{\beta}{2}).$$

6

However, this equality cannot be satisfied: when $l = (1, 1, 0, \cdots 0, \cdots)$, $\sum_{t=0}^{\infty} \beta^t l_t = 1 + \beta = \frac{1-\beta^2}{1-\beta} > \frac{1-\frac{\beta}{2}}{1-\beta}$ and when $l = (1, 0, 1, 1, \cdots)$, we have $\sum_{t=0}^{\infty} \beta^t l_t = 1 + \frac{\beta^2}{1-\beta} = \frac{1+\beta(\beta-1)}{1-\beta} < \frac{1-\frac{\beta}{2}}{1-\beta}$. These are the only two relevant cases and we obtain $V^2 = V^1 > V^0$. Although the sup-inf and the inf-sup FE's have the same value there does not exist a saddle point to the sequential problem and the optimal lottery-free solution cannot be characterized by a FE.

3 A general formulation

We consider an optimization problem with forward looking constraints similar to the problems in Cole and Kubler (2012) or Marcet and Marimon (2019). Time is discrete and infinite, $t = 0, 1, \ldots$. Exogenous shocks (s_t) follow a Markov chain with transition π and realize in a finite set $\mathcal{S}$. A history of shocks up to some time t is denoted by $s^t = (s_0, \ldots, s_t) \in \mathcal{S}^t$. The planner chooses actions contingent on each shock history, $(a(s^t))_{t \in \mathbb{N}, s^t \in \mathcal{S}^t} \in \mathcal{A}^\infty$. At each s^t, the value of the physical state is denoted by $x(s^t)$, and an action, a, is feasible if $p(x(s^t), a(s^t), s_t) \geq 0$ for any $t = 0, 1, \cdots, s^t \in \mathcal{S}^t$. The physical state evolves according to $x(s^t) = \zeta(x(s^{t-1}), a(s^{t-1}), s_{t-1})$ for $t = 1, 2, \ldots$, with $x(s^0) = x_0$. Given γ, x_0, s_0, the optimization problem is as follows.

$$\max_{(a(s^t)) \in \mathcal{A}^\infty \subset \ell^\infty} \mathbb{E}_{s_0, x_0} \sum_{t=0}^{\infty} \beta^t \left(r(x(s^t), a(s^t), s_t) + \sum_{i=1}^{I} \gamma^i g^i(x(s^t), a(s^t), s_t) \right)$$

$$\text{s.t. } \mathbb{E}_{s_t} \sum_{n=0}^{N_i} \beta^n g^i(x(s^{t+n}), a(s^{t+n}), s_{t+n}) \geq \bar{g}^i(x(s^t), a(s^t), s_t), \quad \forall t \in \mathbb{N}, \forall s^t \in \mathcal{S}^t, \forall i \in \{1, \cdots, I\},$$

$$\text{where } x(s^{t+1}) = \zeta(x(s^t), a(s^t), s_t) \text{ and } p(x(s^t), a(s^t), s_t) \geq 0, \quad \forall t \in \mathbb{N}, s^t \in \mathcal{S}^t.$$
$$(3.1)$$

There are $i = 1, \ldots, I \geq 1$ forward looking constraints, and each constraint i restricts actions over the next N_i periods. For notational simplicity, we focus on the case where $N_i = \infty$ for all $i = 1, \ldots, I$. In Appendix E we prove that all our arguments can be trivially extended to the general case. We also simply write $\bar{g}^i$ instead of taking it as a function of the state and actions. Throughout the paper, we make the following assumptions.

Assumption 3.1

1. $(s_t)_{t=0}^{\infty}$ is a Markov process. The set of Markov states $\mathcal{S}$ is a finite set, and $\pi(s'|s) > 0, \forall s, s' \in \mathcal{S}$.

2. $\mathcal{A} \subset \mathbb{R}^n$ is a finite set, $\mathcal{X} \subseteq \mathbb{R}^m$ is a countable set. ζ is map from $\mathcal{X} \times \mathcal{A} \times \mathcal{S}$ to $\mathcal{X}$.

3. The functions $r : \mathcal{X} \times \mathcal{A} \times \mathcal{S} \to \mathbb{R}$, and $g^i : \mathcal{X} \times \mathcal{A} \times \mathcal{S} \to \mathbb{R}, \forall i \in \{1, \cdots, I\}$ are bounded.

4. $0 < \beta < 1$.

5. For any $s_0 \in \mathcal{S}$ and $x_0 \in \mathcal{X}$, there exists a feasible point $(a(s^t)) \in \mathcal{A}^\infty$ to (3.1).

The assumption that $\mathcal{A}$ is finite has no practical relevance since it can be arbitrarily large. This assumption is made for theoretical convenience and it guarantees that the constraints map to ℓ^∞.

At each t, the set of possible pre-action histories up to t is defined by

$$\mathcal{H}^t := \mathcal{S}^{t+1} \times \mathcal{A}^t,$$

where any $h^t \in \mathcal{H}^t$ can be represented as

$$h^t = (s_0, a_0, \cdots, s_{t-1}, a_{t-1}, s_t).$$

7

Note that the history up to time t includes the exogenous shock realized at t but not the action at t.

We define the set of feasible actions given x and s as $\tilde{\mathcal{A}}(x,s) := \{a \in \mathcal{A}|p(x,a,s) \geq 0\}$. For given $x_0 \in \mathcal{X}$, $s_0 \in \mathcal{S}$, we define the set $\tilde{\mathcal{A}}^\infty = \tilde{\mathcal{A}}^\infty(x_0) \subset \mathcal{A}^\infty$ as

$$\tilde{\mathcal{A}}^\infty(x_0) := \{(a_t(s^t)) \in \mathcal{A}^\infty | a(s^t) \in \tilde{\mathcal{A}}(x(s^t), s_t),\, x(s^{t+1}) = \zeta(x(s^t), a(s^t), s_t)\, (\forall t \in \mathbb{N},\, s^t \in \mathcal{S}^t)\},$$

ensuring that the constraints $p(x_t, a_t(s^t), s_t) \geq 0$ hold for all t. We also sometime write $x_t(a^{t-1}) = x(s^t, a^{t-1})$ to emphasize that the physical state at s^t depends on the history of actions up to $t-1$, $a^{t-1} = (a_0, \ldots, a_{t-1})$. In the following, in particular in the following sections and in the appendices, we will often simply write x_t and a_t instead of $x(s^t)$ and $a(s^t)$, whenever there is no scope for confusion. Note that by stationarity $(a(s^t)) \in \tilde{\mathcal{A}}^\infty(x_0)$ if and only if $a(s_0) \in \tilde{\mathcal{A}}(x_0, s_0)$, and for all $t \geq 1$, $a(s^t) \in \tilde{\mathcal{A}}^\infty(\zeta(x_0, a(s_0), s_0))$ for any $s_0 \in \mathcal{S}$.

We define Λ as the ℓ^1 space of the "non-normalized" Lagrangian multipliers $((\beta^t \lambda^i(h^t) \pi^t(s^t|s_0))_{t \in \mathbb{N},\, h^t \in \mathcal{H}^t}$, i.e.

$$(\lambda^i(h^t)) \in \Lambda := \{(\lambda^i(h^t) \geq 0)| \sum_{t=0}^\infty \sum_{h^t \in \mathcal{H}^t} \sum_{i=1}^I \beta^t \lambda^i(h^t) \pi^t(s^t|s_0) < \infty\}. \tag{3.2}$$

We define the Lagrangian functional as

$$L((a(s^t)), (\lambda^i(h^t)); (\gamma^i), x_0, s_0)$$
$$= \mathbb{E}_{s_0, x_0} \sum_{t=0}^\infty \beta^t \left(r(x(s^t, a^{t-1}), a(s^t), s_t) + \sum_{i=1}^I \gamma^i g^i(x(s^t, a^{t-1}), a(s^t), s_t) \right.$$
$$\left. + \sum_{i=1}^I \lambda^i(h^t) \left(\sum_{n=0}^\infty \beta^n g^i(x(s^{t+n}, a^{t+n-1}), a(s^{t+n}), s_{t+n}) - \bar{g}^i \right) \right). \tag{3.3}$$

In this section, we show that the inf-sup problem for the Lagrangian functional can be solved recursively. Having established this, we proceed in Section 4 to prove that the inf-sup value corresponds to a Lagrangian dual of a system that is a minor variation of (3.1) (namely (4.6)) below . Consequently, it must also equal the optimal value of a lottery problem.

3.1 Recursive Formulations

3.1.1 Recursive Dual Value Function – the inf-sup FE

For any $\gamma \in \mathbb{R}_+^I$, $x_0 \in \mathcal{X}$, $s_0 \in \mathcal{S}$, we define the dual value function as

$$D(\gamma, x_0, s_0) = \inf_{(\lambda^i(h^t)) \in \Lambda} \sup_{(a(s^t)) \in \tilde{\mathcal{A}}^\infty(x_0)} L((a(s^t)), (\lambda^i(h^t)); (\gamma^i), x_0, s_0). \tag{3.4}$$

We have the following theorem.

Theorem 3.1 *For any $x \in \mathcal{X}$, $s \in \mathcal{S}$, $\gamma \in \mathbb{R}_+^I$, the dual value function $D(\gamma, x, s)$ defined in* (3.4) *satisfies the following recursive equation*

$$D(\gamma, x, s) = \inf_{\lambda \in \mathbb{R}_+^I} \sup_{a \in \tilde{\mathcal{A}}(x,s)} \left[\left(r(x, a, s) + \sum_{i=1}^I \left(\gamma^i g^i(x, a, s) + \lambda^i(g^i(x, a, s) - \bar{g}^i) \right) \right) + \beta \mathbb{E}_s D(\gamma + \lambda, x', s') \right], \tag{3.5}$$

where $x' = \zeta(x, a, s)$.

The spirit of the proof is the same as in the proof for Proposition 3 in Pavoni et al. (2018a). The proof does not rely on a min-max theorem but instead, it relies only on the following lemma, which allows us to pass each inf operator to the front of the expression.

8

Lemma 3.1 *Suppose that X, Y are subsets of Banach spaces, and $f : X \times Y \to \mathbb{R}$ is an arbitrary functional, $\Gamma := \{\boldsymbol{y} : X \to Y\}$. Suppose that $\tilde{X} \subseteq X$, then*

$$\inf_{\boldsymbol{y} \in \Gamma} \sup_{x \in \tilde{X}} f(x, \boldsymbol{y}(x)) = \sup_{x \in \tilde{X}} \inf_{\boldsymbol{y} \in \Gamma} f(x, \boldsymbol{y}(x)).$$

3.1.2   The sup-inf functional equation

What would happen if one interchanged the inf and sup operator in the inf-sup functional equation (3.5)? Clearly, the optimal value is still related to the dual formulation of the problem (and generally not to the primal). To understand the issue in more detail, we define the post-action history set slightly differently as

$$\tilde{\mathcal{H}}^t = \mathcal{S}^{t+1} \times \mathcal{A}^{t+1},$$

where any $\tilde{h}^t \in \tilde{\mathcal{H}}$ is represented as

$$\tilde{h}^t = (s_0, a_0, \cdots, s_t, a_t), \text{ or } \tilde{h}^t = (s^t, a^t), \text{ where } a^t = (a_0, \cdots, a_t).$$

The difference between h^t and $\tilde{h}^t$ lies in the fact that $\tilde{h}^t$ includes the history until time t including the shock *and* the action at t while h^t only includes the shock (and history until time t).

We allow the Lagrange multiplier associated with the forward looking constraint starting at t to depend on $\tilde{h}^t$ (and not only on h^t), and define $\tilde{\Lambda}$ as the ℓ^1 space of the "non-normalized" Lagrangian multipliers $((\beta^t \lambda^i(\tilde{h}^t) \pi^t(s^t | s_0))_{t \in \mathbb{N}, \tilde{h}^t \in \tilde{\mathcal{H}}^t}$ analogously to (3.2). The Lagrangian function then becomes

$$
\begin{aligned}
&\tilde{L}((a(s^t)), (\lambda^i(\tilde{h}^t)); (\gamma^i), x_0, s_0) \\
=&\mathbb{E}_{s_0} \sum_{t=0}^{\infty} \beta^t \left(r(x(s^t, a^{t-1}), a(s^t), s_t) + \sum_{i=1}^{I} \gamma^i g^i(x(s^t, a^{t-1}), a(s^t), s_t) \right. \\
&+ \left. \sum_{i=1}^{I} \lambda^i(\tilde{h}^t) \left(\sum_{n=0}^{\infty} \beta^n g^i(x(s^{t+n}, a^{t+n-1}), a(s^{t+n}), s_{t+n}) - \bar{g}^i \right) \right).
\end{aligned}
\tag{3.6}
$$

For any $\gamma \in \mathbb{R}_+^I$, $x_0 \in \mathcal{X}$, $s_0 \in \mathcal{S}$, we define the dual value function as

$$\tilde{D}(\gamma, x_0, s_0) = \inf_{(\lambda^i(\tilde{h}^t)) \in \tilde{\Lambda}} \sup_{(a(s^t)) \in \tilde{\mathcal{A}}^{\infty}(x_0)} \tilde{L}((a(s^t)), (\lambda^i(\tilde{h}^t)); (\gamma^i), x_0, s_0). \tag{3.7}$$

The following theorem follows directly from Lemma 3.1 and it motivates our interest in the sup-inf functional equation.

Theorem 3.2 *For any $x \in \mathcal{X}$, $s \in \mathcal{S}$, $\gamma \in \mathbb{R}_+^I$, the dual value function $\tilde{D}(x, \gamma, s_0)$ defined in (3.7) satisfies the following recursive equation*

$$\tilde{D}(\gamma, x, s) = \sup_{a \in \tilde{\mathcal{A}}(x,s)} \inf_{\lambda \in \mathbb{R}_+^I} \left[\left(r(x, a, s) + \sum_{i=1}^{I} \left(\gamma^i g^i(x, a, s) + \lambda^i (g^i(x, a, s) - \bar{g}^i) \right) \right) + \beta \mathbb{E}_s \tilde{D}(\gamma + \lambda, x', s') \right],$$
$$\tag{3.8}$$

where $x' = \zeta(x, a, s)$.

9

3.2 Existence of a solution to the FE

We now establish the existence of solutions to both functional equations. Although both inf-sup and sup-inf Bellman operators exhibit monotonicity and discounting properties, the unbounded nature of Lagrangian multipliers and the resulting unboundedness of the dual value function preclude the direct application of standard contraction arguments from Stokey and Lucas Jr (1989). Bloise and Siconolfi (2022) demonstrate the existence of fixed points for the Negishi operator using Tarski's fixed point theorem[3]. Alternatively, Pavoni et al. (2018a) employed Thompson's metric to establish contraction, requiring the additional technical assumption that there exists some function F satisfying $\mathcal{B}(F) > F + \|\gamma\|\epsilon$. The framework achieves a contraction factor of $1 - O(\epsilon)$ (under Thompson's metric). We develop a constructive proof for the existence of fixed points that directly leverages the monotonicity structure of the Bellman operator. Our method provides an explicit iterative scheme that converges to the solution while avoiding the additional assumptions required in Pavoni et al. (2018a).

Throughout, we want to require the value function to be Lipschitz continuous with respect to the multiplier and we define an upper bound on the Lipschitz constant as

$$L = \frac{\|r\|_\infty + \sum_{i=1}^I \|g_i\|_\infty}{1 - \beta}.$$

Definition 3.1 *For any $k > 0$, we define*

$$B(k) = \{\gamma \in \mathbb{R}_+^I : \|\gamma\|_\infty \le k\}.$$

We define the function space

$$\mathcal{M} := \{F : \mathbb{R}_+^I \times \mathcal{X} \times \mathcal{S} \to \mathbb{R} :$$
$$(i) F(\cdot, x, s) \in L^\infty(B(k)), \ \forall k \in \mathbb{N}_+, x \in \mathcal{X}, \ s \in \mathcal{S};$$
$$(ii) \sum_{i=1}^{|\mathcal{S}|} \frac{1}{2^i} \left[\sum_{j=1}^\infty \frac{1}{2^j} \left(\sum_{k=1}^\infty \frac{1}{2^k} \|F(\cdot, x_j, s_i)\|_{L^\infty(B(k))} \right) \right] < \infty \}.$$

For any $F \in \mathcal{M}$, the norm[4] of F is defined by

$$\|F\|_{\mathcal{M}} := \sum_{i=1}^{|\mathcal{S}|} \frac{1}{2^i} \left[\sum_{j=1}^\infty \frac{1}{2^j} \left(\sum_{k=1}^\infty \frac{1}{2^k} \|F(\cdot, x_j, s_i)\|_{L^\infty(B(k))} \right) \right].$$

It is easy to verify that the space $(\mathcal{M}, \|\cdot\|_{\mathcal{M}})$ is a complete metric space.

---

[3]However, Tarski's fixed point theorem is generally non-constructive when the continuity of the Negishi operator remains unverified. For instance, $f : [0, 1] \to [0, 1]$ defined as

$$f(x) = \begin{cases} \frac{1}{2} + (x - \frac{1}{2})^2, & \frac{1}{2} < x \le 1 \\ x^2, & 0 \le x \le 1/2 \end{cases}$$

is a monotone operator, and the existence of fixed points($x^{FP} = 0$) can be guarented by the Tarski's fixed point theorem. However, $f^{(\infty)}(1) = \frac{1}{2}$ is not a fixed point of f, due to the non-continuity of f at $x = \frac{1}{2}$.

[4]Convergence in this norm implies local uniform convergence.

Definition 3.2 *We define the subset $\mathcal{N} \subset \mathcal{M}$ as follows*

$$\mathcal{N} := \{F \in \mathcal{M} :$$

$$(i)\ F(\cdot, x, s)\ \text{is convex; } (ii) F(\cdot, x, s)\ \text{is L-Lipschitz;} \tag{3.9}$$

$$(iii) F(\gamma, x, s) \geq v^0 + \sum_{i=1}^{I} \gamma^i v^i, \tag{3.10}$$

$$where \begin{cases} v^0 = \mathbb{E}_{s_0=s} \sum_{t=0}^{\infty} \beta^t r(x_t, a_t(s^t), s_t); \\ v^i = \mathbb{E}_{s_0=s} \sum_{t=0}^{\infty} \beta^t g^i(x_t, a_t(s^t), s_t), \quad i \in \{1, \cdots, I\}, \end{cases}$$

$$for\ any\ feasible\ (a_t(s^t))_{t \geq 0}\ to\ problem\ (3.1)\ with\ x_0 = x,\ s_0 = s \tag{3.11}$$

$$(iv) F(\gamma, x, s) \leq (1 + \sum_{i=1}^{I} \gamma^i) L \}. \tag{3.12}$$

Note that (iii) implies that

$$F(\gamma, x, s) \geq -(1 + \sum_{i=1}^{I} \gamma^i) L.$$

It is straightforward to verify that $\mathcal{N}$ is a closed subset of $\mathcal{M}$. We define the Bellman operator as follows.

Definition 3.3 *Let $\mathcal{F}$ denote the set of functions $F : \mathbb{R}_+^I \times \mathcal{X} \times \mathcal{S} \to \mathbb{R}$. The Bellman operator $\mathcal{B} : \mathcal{F} \to \mathcal{F}$ is defined by*

$$\mathcal{B}(F)(\gamma, x, s) := \inf_{\lambda \in \mathbb{R}_+^I} \sup_{a \in \tilde{\mathcal{A}}(x,s)} \left[\left(r(a, x, s) + \sum_{i=1}^{I} \left(\gamma^i g^i(a, x, s) + \lambda^i (g^i(a, x, s) - \bar{g}^i) \right) \right) + \beta \mathbb{E}_s F(\gamma + \lambda, x', s') \right]$$

where $x' = \zeta(x, a, s)$.

The following monotonicity property of $\mathcal{B}$ can be verified directly:

Lemma 3.2 *For $F_1, F_2 \in \mathcal{M}$, s.t. $F_1(x, \gamma, s) \leq F_2(x, \gamma, s)$, $\forall (x, \gamma, s) \in \mathcal{X} \times \mathbb{R}_+^I \times \mathcal{S}$, we have $\mathcal{B}(F_1)(x, \gamma, s) \leq \mathcal{B}(F_2)(x, \gamma, s)$, $\forall (x, \gamma, s) \in \mathcal{X} \times \mathbb{R}_+^I \times \mathcal{S}$.*

Moreover, $\mathcal{B}$ maps $\mathcal{N}$ into itself.

Lemma 3.3 *Under Assumption 3.1, the Bellman operator $\mathcal{B}$ maps $\mathcal{N}$ to itself, that is,*

$$\mathcal{B}(\mathcal{N}) \subset \mathcal{N}.$$

With this we can state the following theorem.

Theorem 3.3 *Assume that $F_0(\gamma, x, s) = \left(1 + \sum_{i=1}^{I} \gamma^i\right) L$. Then there exists $F^* \in \mathcal{N}$, such that*

$$\|\mathcal{B}^{(n)}(F_0) - F^*\|_{\mathcal{M}} \to 0, \ as\ n \to \infty,$$

and F^* is the largest fixed point of $\mathcal{B}$ in $\mathcal{N}$, i.e.

1. *$\mathcal{B}(F^*) = F^*$, and*
2. *For any $F \in \mathcal{N}$ s.t. $\mathcal{B}(F) = F$, we have $F \leq F^*$.*

11

To show that the largest fixed point obtained from Theorem 3.3 equals to the dual value function D, it suffices to verify that $D \in \mathcal{N}$, and that D is not less than any fixed point in $\mathcal{N}$. The property that $D \in \mathcal{N}$ is directly to verified after we show that D equals to the optimal value of the lottery system in the next section. The next lemma shows that D is not less than any fixed point in $\mathcal{N}$.

Lemma 3.4 *Suppose that $F \in \mathcal{N}$ is a fixed point of $\mathcal{B}$. Then $F(x, \gamma, s) \le D(x, \gamma, s)$ for all $x \in \mathcal{X}$, $\gamma \in \mathbb{R}_+^I$, $s \in \mathcal{S}$.*

The same argument can be used for the sup-inf problem to establish the following corollary.

Corollary 3.1 *The following functional equation,*

$$F(\gamma, x, s) = \sup_{a \in \tilde{\mathcal{A}}(x,s)} \inf_{\lambda \in \mathbb{R}_+^I} \left[\left(r(a, x, s) + \sum_{i=1}^I \left(\gamma^i g^i(a, x, s) + \lambda^i (g^i(a, x, s) - \bar{g}^i) \right) \right) + \beta \mathbb{E}_s F(\gamma + \lambda, x', s') \right]$$

where $x' = \zeta(x, a, s)$, has a solution in $\mathcal{N}$. Moreover, every solution $F(.)$ satisfies $F(x, \gamma, s) \le \tilde{D}(x, \gamma, s)$ for all (x, γ, s) .

4 Lotteries

In the case discussed in Marcet and Marimon (2019) where the objective function is concave and the constraint-set is convex, the inf-sup and the sup-inf functional equations have the same value and this is equal to the optimal value of the original problem. As we have seen in our introductory example, under non-convexities they might have different values and their values might be strictly larger than the value of the original problem. In this section we present the main results of the paper that show that both the inf-sup and the sup-inf functional equation have natural economic interpretations if one allows for lotteries. The main results of this paper establish a relation between the sup-inf and the inf-sup FE and a modified problem where the planner has access to lotteries. The theoretical foundation for this relationship lies in the fact that the duality gap can be bridged by randomization. To make this precise we need to introduce a few mathematical facts.

4.1 The duality gap and randomization

To formalize these ideas, we first consider an abstract setting of constrained optimization in Banach-spaces. To be consistent with standard conventions in optimization theory, we consider the minimization problem instead of the maximization one in this section.

Definition 4.1 *Let X be a Banach space, $\Omega \subseteq X$ be an arbitrary subset of X, Y be a Banach space with a closed positive cone $P \subset Y$, $f : \Omega \to \mathbb{R} \cup \{+\infty\}$ be a proper extended real-valued functional, and $g : \Omega \to Y$ be an arbitrary functional. The **optimization problem**, is defined by*

$$\inf_{x \in \Omega} f(x),$$
$$\text{s.t. } g(x) \le \theta_Y, \tag{4.1}$$

where $g(x) \le \theta_Y$ means $g(x) \in -P$.

We define the **perturbation functional** $v : Y \to \mathbb{R} \cup \{+\infty, -\infty\}$ by letting $v(y)$ be the optimal value of the following **perturbed problem**

$$v(y) = \inf_{x \in \Omega} f(x),$$
$$\text{s.t. } g(x) \le y, \tag{4.2}$$

12

where $g(x) \leq y$ means $g(x) - y \in -P$. The **Lagrangian functional** $L : \Omega \times Y_+^* \to \mathbb{R} \cup \{+\infty\}$ to the optimization problem (4.1) is defined by

$$L(x, y^*) = f(x) + \langle y^*, g(x) \rangle. \tag{4.3}$$

Here Y_+^* denotes the set

$$\{y^* \in Y^* | \langle y^*, y \rangle \geq 0, \, \forall y \geq \theta_Y \}.$$

Let X, Ω, Y, f, g be defined as in Definition 4.1. The **inf-sup problem**, or the **primal problem**, is defined by

$$p := \inf_{x \in \Omega} \sup_{y^* \in Y_+^*} L(x, y^*). \tag{4.4}$$

Similarly, the **sup-inf problem**, or the **dual problem**, is defined by

$$d := \sup_{y^* \in Y_+^*} \inf_{x \in \Omega} L(x, y^*). \tag{4.5}$$

The Dual Problem for Lagrangian (3.3)   We consider the modifed problem of (3.1) as follows

$$\sup_{(a(s^t)) \in \tilde{\mathcal{A}}^\infty(x_0) \subset \ell^\infty} \mathbb{E}_{s_0} \sum_{t=0}^\infty \beta^t \left(r(x(s^t), a(s^t), s_t) + \sum_{i=1}^I \gamma^i g^i(x(s^t), a(s^t), s_t) \right),$$

$$\text{s.t. } 1_{\{(\tilde{a}_0, \cdots, \tilde{a}_{t-1}) = (a(s_0), \cdots, a(s^{t-1}))\}} \left(\mathbb{E}_{s_t} \sum_{n=0}^\infty \beta^n g^i(x(s^{t+n}), a(s^{t+n}), s_{t+n}) - \bar{g}^i \right) \geq 0,$$

$$\forall t \in \mathbb{N}, \, \forall h^t = (s_0, \tilde{a}_0, \cdots, s_{t-1}, \tilde{a}_{t-1}, s_t) \in \mathcal{H}^t, \, \forall i \in \{1, \cdots, I\}. \tag{4.6}$$

For this problem, we define the functional

$$f : \tilde{\mathcal{A}}^\infty(x_0) \to \mathbb{R}, \quad (a(s^t)) \mapsto \mathbb{E}_{s_0} \sum_{t=0}^\infty \beta^t \left(r(x(s^t), a(s^t), s_t) + \sum_{i=1}^I \gamma^i g^i(x(s^t), a(s^t), s_t) \right), \tag{4.7}$$

and the map

$$g : \tilde{\mathcal{A}}^\infty(x_0) \to \ell^\infty,$$
$$(a(s^t)) \mapsto (g_{t,h^t,i}(a(s^t)))_{t \in \mathbb{N}, \, h^t \in \mathcal{H}^t, \, i \in \{1, \cdots, I\}} :=$$
$$\left(1_{\{(\tilde{a}_0, \cdots, \tilde{a}_{t-1}) = (a(s_0), \cdots, a(s^{t-1}))\}} \left(\mathbb{E}_{s_t} \sum_{n=0}^\infty \beta^n g^i(x(s^{t+n}), a(s^{t+n}), s_{t+n}) - \bar{g}^i \right) \right)_{t \in \mathbb{N}, \, h^t \in \mathcal{H}^t, \, i \in \{1, \cdots, I\}}. \tag{4.8}$$

Problem (4.6) can then be formulated as

$$\sup_{(a(s^t)) \in \tilde{\mathcal{A}}^\infty(x_0) \subset \mathcal{A}^\infty} f((a(s^t))), \tag{4.9}$$
$$\text{s.t. } g((a(s^t))) \geq 0.$$

Therefore, according to the definition of the dual problem (B.4), the dual problem of (C.13) is defined as[5]

$$d = \inf_{\lambda \in \ell_+^{\infty,*}} \sup_{(a(s^t)) \in \tilde{\mathcal{A}}^\infty(x_0)} f((a(s^t))) + \langle \lambda, g((a(s^t))) \rangle. \tag{4.10}$$

---

[5]Note that the primal problem now is a maximization problem, and hence the dual problem has the inf-sup form.

To establish the existence of Lagrange multipliers in the ℓ^1 space—rather than the $\ell^{\infty,*}$ space considered in (4.10)—and to consequently show that $D(\gamma, x_0, s_0)$ defined in (3.4) equals the dual value d from (B.3), we follow the approach of Dechert (1982) and Pavoni et al. (2018a). The detailed argument is provided in Appendix C (see Theorem C.3).

In general $d \leq p$ and the difference between the two values is called the duality-gap. To show how this gap can be bridged by convexification, we need the following definitions.

Definition 4.2 *Let $f : X \to \mathbb{R} \cup \{+\infty, -\infty\}$ be an extended real-valued functional. Its epigraph is defined as*

$$epi(f) := \{(x, r) \in X \times \mathbb{R} | f(x) \leq r\}.$$

The functional $f^* : X^* \to \mathbb{R} \cup \{+\infty, -\infty\}$ defined by

$$f^*(x^*) := \sup_{x \in X} \{\langle x^*, x \rangle - f(x)\}$$

*is called the **convex conjugate**, or **conjugate** of f, and the mapping*

$$f \mapsto f^*$$

*is called the **Legendre-Fenchel** transformation. Furthermore, $f^{**} : X \to \mathbb{R} \cup \{+\infty, -\infty\}$[6] defined by*

$$f^{**}(x) := \sup_{x^* \in X} \{\langle x^*, x \rangle - f^*(x^*)\}$$

*is called the **biconjugate** of f.*

The following theorem which is a variation of the Fenchel-Moreau theorem constitutes the theoretical foundation of our analysis. It seems to be known in mathematics[7] but a clean reference is difficult to find. For completeness, we provide a proof in Appendix A.

Theorem 4.1 *Let X be a Banach space, $f : X \to \mathbb{R} \cup \{+\infty\}$ be a proper extended real-valued functional and assume that there exists $\underline{x}^* \in X^*$, $\underline{\beta} \in \mathbb{R}$, such that*

$$\langle \underline{x}^*, x \rangle + \underline{\beta} \leq f(x), \quad \forall x \in X. \tag{4.11}$$

Then the following holds.

$$epi(f^{**}) = cl \ co \ epi(f),$$

where cl co denotes the closure of the convex hull.

The following theorem states that this duality gap can be bridged if randomization is possible in the sense that the epigraph of the perturbed dual problem is the closure of the convex hull of the epigraph of the perturbed primal problem.

Theorem 4.2 *Let X, Ω, Y, f, g be defined as in Definition 4.1. Then $p = v(\theta_Y)$, $d = v^{**}(\theta_Y)$. where $v^{**}(.)$ denotes the biconjugate of v. Moreover*

$$epi(v^{**}) = cl \ co \ epi(v).$$

The basic intuition for the results in a finite dimensional setting is explained in detail in Theorems 2.3 and 2.4 in Shen et al. (2025).

---

[6]If X is reflexive, then $X = X^{**}$, hence f^{**} is indeed the conjugate of f^*. If X is not reflexive, then $X \subset X^{**}$ is a closed subset of X^{**}, hence f^{**} can be regarded as the conjugate of f^* restricted in X.

[7]According to theorem 5 in Rockafellar (1974)(page 16), one might obtain the same conclusion when f satisfies lsc co $(f) > -\infty$, $\forall x \in X$, which is a weaker condition than ours. However, the proof in Rockafellar (1974) relies on a geometric result that was not rigorously stated in the book. The finite-dimensional version of this theorem was provided in Theorem 1.3.5 in Urruty and Lemaréchal (1993)(page 45).

4.2 Lotteries and the dual problem

Motivated by Theorem 4.2, we now want to explore how the interpretation and the optimal solution of (3.1) change if the planner can randomize of actions. Since $\mathcal{A}$ is compact, according to the Tychonoff theorem, $\mathcal{A}^\infty$ is compact in the product topology. Furthermore, since the topology in $\mathcal{A}$ is metrizable, the product topology in $\mathcal{A}^\infty$ is metrizable(see Theorem 3.36 in Aliprantis and Border (2006)). Given $x_0 \in \mathcal{X}$, we know that $\tilde{\mathcal{A}}^\infty(x_0) \subseteq \mathcal{A}^\infty$ and is a closed subset in $\mathcal{A}^\infty$ in the product topology. Hence $\tilde{\mathcal{A}}^\infty(x_0)$ is compact in the product topology. Since $\mathcal{A}$ is finite, and that the discount factor satisfies $0 < \beta < 1$, it is straightforward to verify that f and $g_{t,h^t,i}$ defined in (4.7) and (4.8) is continuous under the product topology.

Let $\mathcal{P}(\tilde{\mathcal{A}}^\infty(x_0))$ denote the space of probability measures on $\tilde{\mathcal{A}}^\infty(x_0)$. It is standard to show that $\mathcal{P}(\tilde{\mathcal{A}}^\infty(x_0))$ is compact in the *-weak topology, using Portmanteau Theorem(for closedness, see Theorem 2.1(iv) in Billingsley (2013)) and Prokhorov's theorem(for relative compactness, see Theorem 6.1 in Billingsley (2013)).

Definition 4.3 *The ex-ante lottery problem of* (3.1) *is as follows.*

$$
\sup_{P \in \mathcal{P}(\tilde{\mathcal{A}}^\infty(x_0))} \mathbb{E}_{s_0}^{(a(s^t))\sim P} \sum_{t=0}^\infty \beta^t \left(r(x(s^t, a^{t-1}), a(s^t), s_t) + \sum_{i=1}^I \gamma^i g^i(x(s^t, a^{t-1}), a(s^t), s_t) \right),
$$

$$
\textbf{\textit{s.t.}} \ \ \mathbb{E}_{s^t}^{(a(s^t))\sim P} 1_{\{\tilde{a}^{t-1}=a^{t-1}\}} \left(\sum_{n=0}^\infty \beta^n g^i(x(s^{t+n}, (\tilde{a}^{t-1}, a_t^{t+n-1})), a(s^{t+n}), s_{t+n}) - \bar{g}^i \right) \geq 0,
$$

$$
\forall t \in \mathbb{N}, \ \forall h^t = (s_0, \tilde{a}_0, \cdots, s_{t-1}, \tilde{a}_{t-1}, s_t) \in \mathcal{H}^t, \ \forall i \in \{1, \cdots, I\}.
$$

$$\tag{4.12}$$

In this definition $1_{\{\tilde{a}^{t-1}=a^{t-1}\}}$ denotes the indicator function that is equal to one if the actions realized in the past are equal to $\tilde{a}^{t-1}$ (and zero otherwise) and $a_t^{t+n-1} = (a_t, \ldots, a_{t+n-1})$ denotes the future actions beginning at t up to $t+n-1$ (that are randomized over by the lottery P), with the convention that for $n = 0$, the term is to be ignored. Randomization is performed over infinite sequences of actions $(a(s^t))$ and each forward looking constraint is assumed to hold on average, conditional on the entire history of actions and shocks up to t.

The next theorem shows that, the maximum of problem (4.12) can be achieved because of the *-weak compactness of $\mathcal{P}(\tilde{\mathcal{A}}^\infty(x_0))$, and hence we can replace 'sup' to 'max' in (4.12).

Theorem 4.3 *There exists a maximizer* $P^* \in \mathcal{P}(\tilde{\mathcal{A}}^\infty(x_0))$ *to* (4.12).

It turns out that attention can be limited to lotteries over $\mathcal{A}$ after each history h^t. For this we denote by $\Pi_{h^t \in \mathcal{H}^t} \mathcal{P}(\tilde{A}(h^t))$ the set of sequences of probability measures which satisfy at each history $h^t = (s_0, \tilde{a}_0, \cdots, s_{t-1}, \tilde{a}_{t-1}, s_t) \in \mathcal{H}^t$,

$$
\psi(h^t) \in \mathcal{P}(\tilde{A}(h^t)) := \mathcal{P}(\tilde{A}(x(s^t, \tilde{a}^{t-1}), s_t)).
$$

That is to say, the constraint $p(x_t, a, s_t) \geq 0$ has to hold for each a in the support of $\psi(h^t)$. $a(h^t) \sim \psi(h^t)$ then means $p(a(h^t) = a | a_0 = \tilde{a}_0, \cdots, a_{t-1} = \tilde{a}_{t-1}) = \psi(h^t)(a)$.

The following theorem formalizes that it is sufficient to focus on lotteries in $\Pi_{h^t \in \mathcal{H}^t} \mathcal{P}(\tilde{A}(h^t))$.

Theorem 4.4 *The lottery problem* (4.12) *is equivalent to*

$$\max_{\psi \in \Pi_{h^t \in \mathcal{H}^t} \mathcal{P}(\tilde{A}(h^t))} \mathbb{E}_{s_0}^{(a(h^t)) \sim \psi(h^t))} \sum_{t=0}^{\infty} \beta^t \left(r(x(s^t, a(h^{t-1})), a(h^t), s_t) + \sum_{i=1}^{I} \gamma^i g^i(x(s^t, a(h^{t-1})), a(h^t), s_t) \right),$$

$$s.t. \ \mathbb{E}_{s^t}^{(a(h^t)) \sim \psi(h^t))} 1_{\{\tilde{a}^{t-1} = a^{t-1}\}} \left(\sum_{n=0}^{\infty} \beta^n \left(g^i(x(s^{t+n}, (\tilde{a}^{t-1}, a_t^{t+n-1})), a(s^{t+n}), s_{t+n}) - \bar{g}^i \right) \right) \geq 0,$$

$$\forall t \in \mathbb{N}, \forall h^t = (s_0, \tilde{a}_0, \cdots, s_{t-1}, \tilde{a}_{t-1}, s_t) \in \mathcal{H}^t, \forall i \in \{1, \cdots, I\}.$$

(4.13)

The following theorem states that the value of the ex ante lottery problem is identical to the value of the inf-sup FE. This is the main result of this section that justifies the use of the inf-sup FE in non-convex problem.

Theorem 4.5 *Under Assumption 3.1, the maximum of problem* (4.12)*, denoted as* $V(\gamma, x_0, s_0)$*, equals to* $D(\gamma, x_0, s_0)$*, where* D *is the dual value function of the deterministic problem* (3.1)*.*

As we have demonstrated above, the following corollary provides the fact that $D \in \mathcal{N}$. This result, together with Lemma 3.4, yields that D is largest fixed point of $\mathcal{B}$ in the space $\mathcal{N}$.

Corollary 4.1 *The dual value function satisfies* $D \in \mathcal{N}$*. Therefore, it is equal to the largest fixed point in* $\mathcal{N}$ *of the Bellman operator from Definition 3.3.*

4.3 Ex post lotteries

Alternatively, we can define an ex-post lottery problem that is equivalent to the sup-inf problem (3.8).

Definition 4.4 *The ex-post lottery problem is as follows*

$$\max_{P \in \mathcal{P}(\tilde{A}^{\infty}(x_0))} \mathbb{E}_{s_0}^{(a(s^t)) \sim P} \sum_{t=0}^{\infty} \beta^t \left(r(x(s^t, a^{t-1}), a(s^t), s_t) + \sum_{i=1}^{I} \gamma^i g^i(x(s^t, a^{t-1}), a(s^t), s_t) \right),$$

$$s.t. \ \mathbb{E}_{s^t}^{(a_t(s^t)) \sim P} 1_{\{\tilde{a}^t = a^t\}} \left(\sum_{n=0}^{\infty} \beta^n g^i(x(s^{t+n}, (\tilde{a}^{t-1}, a_t^{t+n-1})), a(s^{t+n}), s_{t+n}) - \bar{g}^i \right) \geq 0,$$

$$\forall t \in \mathbb{N}, \forall \tilde{h}^t = (s_0, \tilde{a}_0, \cdots, s_{t-1}, \tilde{a}_{t-1}, s_t, \tilde{a}_t) \in \tilde{\mathcal{H}}^t, \forall i \in \{1, \cdots, I\}.$$

(4.14)

As in the ex-ante problem, the planner chooses probability distribution over infinite sequences. However, as in our example in Section 2 the difference between the two problems consists of the fact that in the ex post problem the constraint has to hold conditional on each a_t in the support of the lottery. A history is defined as $\tilde{h}^t$ and is assumed to include the action at t. As in the case of ex-ante lotteries, it is equivalent of considering lotteries over actions after each history.

More importantly, We have the following analog of Theorem 4.5.

Theorem 4.6 *Under Assumption 3.1, the maximum of problem* (4.14)*, denoted as* V*, equals to* $\tilde{D}(\gamma, x_0, s_0)$*, where* $\tilde{D}$ *is the value function of the sup-inf FE* (3.8)*. Moreover the largest solution of the FE in* $\mathcal{N}$ *solves* (4.14)*.*

16

4.4 Economic relevance of lotteries

In our simple example in Section 2 the difference between ex ante and ex post lotteries was simply a question of timing. In general principal agent problems with lotteries it is typically assumed that the participation constraint has to hold ex ante while the IC constraint only holds ex post. In this case ex post lotteries would be more relevant (see also Cole and Kubler (2012) for a simple formulation, or Prescott (1999) for a general argument in the static case).

For models of optimal government policy constraints often consist of agents' first order condition. To illustrate the difference between ex post and ex ante lotteries it is useful to consider the simplest possible version of the problem in Aiyagari et al. (2002) and show that ex post lotteries might improve the value in the Ramsey problem. Government expenditures $(g_t)_{t=0}^{\infty}$ are assumed to be exogenous and to follow a finite Markov chain. A single agent takes the risk-free rates (R_t), and income taxes (τ_t), as given, and solves

$$
\max_{(c_t, \ell_t, b_t)_{t=0}^{\infty}} E \sum_{t=0}^{\infty} \beta^t \left(u(c_t) + v(\ell_t) \right)
$$
$$
\text{s.t. } c_t = \ell_t(1 - \tau_t) + b_{t-1} - \frac{b_t}{R_t} \tag{4.15}
$$
$$
\sup_t \|b_t\| < \infty, \ l_t \in [0,1] \forall t
$$

Under full commitment, the government chooses labor income taxes τ_t and interest rates on its debt R_t so that given the optimal choices of the agent, the government budget constraint is satisfied, i.e. for all t, $b_t = \tau_t n_t - g_t + \frac{b_{t+1}}{R_t}$, and that there are no other taxes and interest rates that result in higher utiltiy for the agent. Under the assumption that the first order conditions of the agent's problem are necessary and sufficient, following Marcet and Marimon (2019), we can rewrite the problem as

$$
\max_{(c_t, \ell_t, b_t)_{t=0}^{\infty}} E \sum_{t=0}^{\infty} \beta^t \left(u(c_t) + v(\ell_t) \right)
$$
$$
\text{s.t. } \beta b_t \mathbb{E} u'(c_{t+1}) = u'(c_t)(b_t - c_t) - v'(\ell_t)\ell_t \tag{4.16}
$$
$$
0 \le c_t \le \ell_t - g_t
$$

Note that the problem is slightly different than our general problem since forward looking constraints extend only over two periods. We show in Appendix E how the analysis can be applied to this setting.

Clearly the set of feasible actions is generally not a convex set. The question is if lotteries can improve on the lottery-free solution. Assuming ex-ante lotteries would give a constraint of the form
$$
E^{P(a_t, a_{t+1})} \left(q_t(a_t) u'(c_t(a_t)) + \beta E_t u'(c_{t+1}(a_{t+1})) \right) = 0.
$$
It is difficult to give an economically meaningful interpretation of this constraint. Ex-post lotteries, on the other hand, would yield a constraint of the form

$$
E^{P(a_{t+1})} \left(q_t(a_t) u'(c_t(a_t)) + \beta E_t u'(c_{t+1}(a_{t+1})) \right) = q_t(a_t) u'(c_t(a_t)) + \beta E^{P(a_{t+1})} E_t u'(c_{t+1}(a_{t+1})) = 0,
$$

which has the straightforward interpretation that there is randomness in next period's consumption in addition to the one induced by the exogenous shock.

The case of ex-post lotteries corresponds to the functional equation

$$
W(b, \mu, g) = \sup_{c, b'} \inf_{\gamma \ge 0} \left[u(c) + v(c + g) + \mu b u'(c) + \gamma \left(u'(c)(c - b) + (c + g) v'(c + g) \right) + \beta E_g W(b', \gamma, g') \right].
$$

It makes a big difference both in terms of the optimal value and in terms of the economic interpretation whether one considers the sup-inf FE or the inf-sup FE. The inf-sup functional equation lacks an economic interpretation.

5 Policies

Having established in the previous section that both the inf-sup and sup-inf functional equations yield the correct optimal value for their respective lottery problems, the remaining question becomes how to recover the corresponding optimal lotteries. This section develops an approach to address this problem and determine optimal lottery policies. To build intuition, we begin with the inf-sup equation (3.5). Solving for the value function D (the largest fixed point) also yields, for each state, a minimizer[8] λ^* and a corresponding set of maximizers $a(\lambda^*)$. Although this arises from a problem without lotteries, it provides the key to recovering optimal randomized policies. It follows from Shen et al. (2025) that to ensure that a statewise lottery attains the optimal utility, its support must lie in the set of optimizers, $a(\lambda^*)$. The remaining challenge is to construct an optimal lottery over these support points, with the critical requirement that it satisfies all forward-looking constraints. To this end, we follow the work of Cole and Kubler (2012) and Marimon and Werner (2021) by incorporating the *promised value* into the functional equation to keep track of the forward looking constraints. Based on these insights, we then adapt the iterative method from Shen et al. (2025) to recover the optimal policy.

A key prerequisite for recovering the policy is obviously to establish that the infimum and supremum in (3.5) are attainable. Since the action set is assumed to be finite, this is equivalent to showing that there exists a λ^* that solves (3.5). To this end, we first introduce a variant of Slater's condition.

Assumption 5.1 *There exists $\epsilon > 0$, s.t. for any $s_0 \in \mathcal{S}$, $x_0 \in \mathcal{X}$, there exists a feasible point $(\psi(h^t)) \in \Pi_{h^t \in \mathcal{H}^t} \mathcal{P}(\tilde{A}(h^t))$ to (4.13) satisfying for all $i = 1, \dots I$,*

$$\mathbb{E}_{s_0}^{(a(h^t) \sim \psi(h^t))} \left(\sum_{t=0}^{\infty} \beta^t g^i(x_t, a_t, s_t) - \bar{g}^i \right) \geq \epsilon.$$

Theorem 5.1 *Under Assumption 5.1, for every $x \in \mathcal{X}$, $s \in \mathcal{S}$, $\gamma \in \mathbb{R}_+^I$, there exists λ^* that solves the inf-sup problem (3.5), with D given as the dual value function.*

We introduce a utility promise ϕ^i for $i = 1, \dots I$ and define

$$
\begin{aligned}
W(\gamma_0, x_0, s_0, \phi) = \max_{\psi \in \Pi_{h^t \in \mathcal{H}^t} \mathcal{P}(\tilde{A}(h^t))} & \mathbb{E}_{s_0}^{(a(h^t) \sim \psi(h^t))} \sum_{t=0}^{\infty} \beta^t \left(r(x_t, a_t, s_t) + \sum_{i=1}^{I} \gamma^i g^i(x_t, a_t, s_t) \right), \\
\text{s.t. } & \mathbb{E}_{h^t}^{(a(h^t) \sim \psi(h^t))} \left(\sum_{n=0}^{\infty} \beta^n g^i(x_{t+n}, a_{t+n}, s_{t+n}) - \bar{g}^i \right) \geq 0, \\
& \forall t \in \mathbb{N}, \forall h^t = (s_0, \tilde{a}_0, \cdots, s_{t-1}, \tilde{a}_{t-1}, s_t) \in \mathcal{H}^t, \forall i \in \{1, \cdots, I\}; \\
& \mathbb{E}_{s_0}^{(a(h^t) \sim \psi(h^t))} \left(\sum_{t=0}^{\infty} \beta^t g^i(x_t, a_t, s_t) - \phi^i \right) \geq 0.
\end{aligned}
$$
(5.1)

In general, we have $W(\gamma, x, s, \phi) \leq D(\gamma, x, s)$, because we introduce an additional constraint for W. The following lemma shows the equivalence between the feasible promised values and the subdifferential of D which we denote by $\partial D(\gamma, x, s)$.

---

[8]The minimizer might not be unique but the argument holds for any minimizer.

Lemma 5.1 *For any $\gamma \in \mathbb{R}_+^I$, $x \in \mathcal{X}$, $s \in \mathcal{S}$, the following are equivalent:*

1. $\phi \in \partial D(\gamma, x, s)$;

2. *There exists $P^* \in \mathcal{P}(\tilde{\mathcal{A}}^\infty(x_0))$, s.t.*

   - P^* *maximizes problem* (4.12), *and*

   - $\mathbb{E}_{s_0}^{(a_t \sim P^*)} \sum_{t=0}^\infty \beta^t g(x_t, a_t, s_t) = \phi$.

The following is a direct corollary of this lemma and gives conditions that ensure that $W(\gamma, x, s, \phi) = D(\gamma, x, s)$.

Corollary 5.1 *For any $\gamma \in \mathbb{R}_+^I$, $x \in \mathcal{X}$, $s \in \mathcal{S}$, $W(\gamma, x, s, \phi) = D(\gamma, x, s)$ if and only if there exists $\phi_D \in \partial D(\gamma, x, s)$, s.t. $\phi \leq \phi_D$.*

To derive a functional equation for W, we follow the proof of Theorem 3.1 and combining it with the result of Corollary 5.1, we obtain the following theorem.

Theorem 5.2 *For any $x \in \mathcal{X}$, $s \in \mathcal{S}$, $\gamma \in \mathbb{R}_+^I$, $\phi \in \mathbb{R}^I$, the dual value function $W(x, \gamma, s_0, \phi)$ defined in* (3.4) *satisfies the following equation*

$$W(\gamma, x, s, \phi) = \inf_{\mu \in \mathbb{R}_+^I, \, \lambda \in \mathbb{R}_+^I} \sup_{a \in \tilde{\mathcal{A}}(x,s)}$$

$$\left[\left(r(x, a, s) + \sum_{i=1}^I \left(\gamma^i g^i(x, a, s) + \lambda^i (g^i(x, a, s) - \bar{g}^i) + \mu^i (g^i(x, a, s) - \phi^i) \right) \right) + \beta \mathbb{E}_s D(\gamma + \lambda + \mu, x', s') \right],$$

where $x' = \zeta(x, a, s)$.

$$(5.2)$$

or, equivalently,

$$W(\gamma, x, s, \phi) = \inf_{\mu \in \mathbb{R}_+^I, \, \lambda \in \mathbb{R}_+^I} \sup_{\psi \in \mathcal{P}(\tilde{\mathcal{A}}(x,s)), \phi'(a,s) \in \mathbb{R}^I} \sum_{a \in \mathcal{A}} \psi(a)$$

$$\left[\left(r(x, a, s) + \sum_{i=1}^I \left(\gamma^i g^i(x, a, s) + \lambda^i (g^i(x, a, s) - \bar{g}^i) + \mu^i (g^i(x, a, s) - \phi^i) \right) \right) + \beta \mathbb{E}_s W(\gamma + \lambda + \mu, x', s', \phi'(a, s')) \right],$$

where $x' = \zeta(x, a, s)$.

$$(5.3)$$

It is easy to verify that if λ^* is a solution to (3.5), then λ^* together with $\mu^* = 0$ is a solution to (5.2) and a solution to (5.3). The following two theorems now describe how an optimal policy can be recovered.

Theorem 5.3 *For any $x \in \mathcal{X}$, $s \in \mathcal{S}$, $\gamma \in \mathbb{R}_+^I$, $\phi \in \partial D(\gamma, x, s)$, the recursive equation* (5.3) *admits at least a solution $(\mu^*, \lambda^*, \psi^*, \phi'^*)$, satisfying*

$$\phi'^*(a, s') \in \partial D(\gamma + \mu^* + \lambda^*, \zeta(x, a, s), s'), \quad \forall a \in supp(\psi^*), s' \in \mathcal{S}, \tag{5.4}$$

$$\left(\sum_{a \in \mathcal{A}} \psi(a)(g(x, a, s) + \beta \mathbb{E}_s \phi'(a, s')) - \phi \geq 0 \right) \perp (\gamma + \mu^*) \tag{5.5}$$

and

$$\left(\sum_{a \in \mathcal{A}} \psi(a)(g(x, a, s) + \beta \mathbb{E}_s \phi'(a, s')) - \bar{g} \geq 0 \right) \perp \lambda^*. \tag{5.6}$$

Using the solutions (ψ^*, ϕ'^*) from Theorem (5.3), we can recover the policy as follows.

19

Theorem 5.4 *Given x_0, s_0, γ_0, $\phi_0 \in \partial D(\gamma_0, x_0, s_0)$. For any $t \geq 0$, $h^t \in \mathcal{H}^t$, we define*

$$(\mu^*(h^t), \lambda^*(h^t), \psi^*(h^t), \phi^*(h^t; a_t, s_{t+1}))$$

a solution to (5.3) *with* $x = x(h^t), s = s_t, \gamma = \gamma(h^{t-1}), \phi = \phi(h^t)$ *satisfying* (5.4)*,* (5.5) *and* (5.6)*, and define* $\gamma^*(h^t) = \gamma^*(h^{t-1}) + \lambda^*(h^t) + \mu^*(h^t)$*. Then* $(\psi^*(h^t))$ *is a solution to* (4.13)*.*

Now the problem reduces to finding a solution to (5.3) satisfying (5.4), (5.5) and (5.6). The following theorem allows us to identify such solutions that bypass the need to consider γ.

Theorem 5.5 *For every* $x \in \mathcal{X}, s \in \mathcal{S}, \gamma \in \mathbb{R}_+^I, \phi \in \partial D(\gamma, x, s)$*, there exists a solution* $(\mu^*, \lambda^*, \psi^*, \phi'^*)$ *to the following problem*

$$W(0, x, s, \phi) = \inf_{\mu \in \mathbb{R}_+^I, \lambda \in \mathbb{R}_+^I} \sup_{\psi \in \mathcal{P}(\tilde{\mathcal{A}}(x,s)), \phi'(a,s') \in \mathbb{R}^I} \sum_a \psi(a)$$

$$\left[\left(r(x, a, s) + \sum_{i=1}^I \left(\lambda^i(g^i(x, a, s) - \bar{g}^i) + \mu^i(g^i(x, a, s) - \phi^i) \right) \right) + \beta \mathbb{E}_s W(\lambda + \mu, x', s', \phi'(a, s')) \right],$$

where $x' = \zeta(x, a, s)$,

(5.7)

that satisfies

$$\phi'^*(a, s') \in \partial D(\lambda^* + \mu^*, x', s'),$$ (5.8)

$$\left(\sum_{a \in \mathcal{A}} \psi^*(a)(g(x, a, s) + \beta \mathbb{E}_s \phi'^*(a, s')) - \phi \geq 0 \right) \perp \mu^*,$$ (5.9)

and

$$\left(\sum_{a \in \mathcal{A}} \psi^*(a)(g(x, a, s) + \beta \mathbb{E}_s \phi'^*(a, s') - \bar{g}) \geq 0 \right) \perp \lambda^*.$$ (5.10)

Moreover, for $(\psi^*, \phi^{*\prime})$*, there exists* $(\tilde{\lambda}^*, \tilde{\mu}^*)$ *such that*

$$(\tilde{\mu}^*, \tilde{\lambda}^*, \psi^*, \phi'^*)$$

is a solution to (5.3) *that satisfies* (5.4)*,*(5.5) *and* (5.6)*.*

Now, it suffices to find a solution to the problem in Theorem 5.5. We introduce a simple algorithm for doing so. We assume that the optimal dual value function $D(.)$ is given. For $\gamma = 0$, $x \in \mathcal{X}$, $s \in \mathcal{S}$, ϕ, we define the inner function in (5.2) as

$$F(\mu, \lambda) := \sup_{a \in \mathcal{A}} \left[\left(r(x, a, s) + \sum_{i=1}^I \left(\lambda^i(g^i(x, a, s) - \bar{g}^i) + \mu^i(g^i(x, a, s) - \phi^i) \right) \right) + \beta \mathbb{E}_s D(\lambda + \mu, x', s') \right].$$

It is easy to check that $F(\mu, \lambda)$ is a convex function and given μ, λ, any

$$\left((g(x, a^*, s) - \phi) + \beta \mathbb{E}_s \partial D(\lambda + \mu, x', s'), (g(x, a^*, s) - \bar{g}) + \beta \mathbb{E}_s \partial D(\lambda + \mu, x', s') \right)$$

is a subgradient of F[9], where $a^* \in \arg\max F$, and $x' = \zeta(x, a^*, s)$.

---

[9]In practical computations, we would use

$$\tilde{\phi} = \left(\frac{D(\lambda + \mu + \epsilon e_i, x', s') - D(\lambda + \mu, x', s')}{\epsilon} \right)$$

to approximately obtain an element in $\partial D(\lambda + \mu, x', s')$.

This observation directly leads to the sub-gradient descend method in Algorithm 1. For a given value of the (extended) state, ϕ, x, s, the method computes an optimal lottery over current actions as well as next periods' utility promises $\phi(s')$ for all $s' \in \mathcal{S}$. Along a simulated path the algorithm can then be used to compute an optimal policy: starting from an unbinding initial promised value[10], we apply Algorithm 1 iteratively at each encountered state to compute the optimal lottery and next period's promised values, which in turn become the input for the subsequent step. This process generates a sequence of action lotteries and promises that defines the optimal policy.

Algorithm 1: Algorithm for the Computation of Optimal Policies

---

Given ϕ, x, s, initial Lagrangian multipliers λ^1, μ^1, the learning rates $\{\sigma^k\}_{k=0}^{\infty}$ and the number of iteration $N > 0$.

For $k = 1 : N$

1. Solve the optimization problem

$$a^k \in \arg \max_{a \in \tilde{\mathcal{A}}(x,s)} \left[\left(r(x,a,s) + \sum_{i=1}^{I} \lambda^{k,i}(g^i(x,a,s) - \bar{g}^i) + \mu^{k,i}(g^i(x,a,s) - \phi^i) \right) + \beta \mathbb{E}_s D(\lambda^k + \mu^k, x', s') \right],$$

and compute

$$\phi^k(s') \in \partial D(\lambda^k + \mu^k, x', s'), \text{ for all } s' \in \mathcal{S}.$$

2. Update the Lagrangian multipliers by

$$\mu^{k+1} = \max\{\mu^k - \sigma^k \left(g(x, a^k, s) - \phi + \beta \mathbb{E}_s \phi^k(s') \right), 0\}$$

$$\lambda^{k+1} = \max\{\lambda^k - \sigma^k \left(g(x, a^k, s) - \bar{g} + \beta \mathbb{E}_s \phi^k(s') \right), 0\}$$

End

1. Construct the stage lottery ψ as

$$\psi^N(a) = \frac{\sum_{k=1}^{N} 1_{a^k=a} \sigma^k}{\sum_{k=1}^{N} \sigma^k}.$$

2. Construct the promised value ϕ' as

$$\phi'^N(a, s') = \frac{\sum_{k=1}^{N} 1_{a^k=a} \sigma^k \phi^k(s')}{\sum_{k=1}^{N} 1_{a^k=a} \sigma^k}.$$

The final theorem of this section formalizes the convergence of this algorithm.

Theorem 5.6 *We consider the Algorithm 1. Assume that the learning rate $\{\sigma^k\}_{k=0}^{\infty}$ satisfies*

$$\sum_{k=0}^{\infty} \sigma^k = \infty, \text{ and } \sum_{k=0}^{\infty} (\sigma^k)^2 < \infty.$$

---

[10]This is equivalent to starting without any promised value, since agents do not need to be promised anything at period 0.

Then $(\lambda^k, \mu^k) \to (\lambda^*, \mu^*)$, for some $(\lambda^*, \mu^*) \in \arg\min F(\lambda, \mu)$. Furthermore, for any $\epsilon > 0$, there exists $\bar{N} > 0$, such that when $N > \bar{N}$, we have

1. $\|(\lambda^k, \mu^k) - (\lambda^*, \mu^*)\| \le \epsilon$,

2. $\psi^N(a)\mathrm{dist}(\phi'^N(a, s'), \partial D(\lambda^* + \mu^*, x', s')) < \epsilon$, for any $a \in \tilde{\mathcal{A}}(x, s)$

3. $\sum_a \psi^N(a)(g(x, a, s) + \beta\mathbb{E}_s\phi'^N(a, s')) - \phi \ge -\epsilon$,

4. $\sum_a \psi^N(a)(g(x, a, s) + \beta\mathbb{E}_s\phi'^N(a, s') - \bar{g}) \ge -\epsilon$,

5. $|\langle \mu^*, \sum_a \psi^N(a)(g(x, a, s) + \beta\mathbb{E}_s\phi'^N(a, s')) - \phi\rangle| \le \epsilon$,

6. $|\langle \lambda^*, \sum_a \psi^N(a)(g(x, a, s) + \beta\mathbb{E}_s\phi'^N(a, s') - \bar{g})\rangle| \le \epsilon$.

Note that for the sup-inf FE the policy can be recovered by the exact same algorithm. By the same arguments as above, we can formulate a functional equation that depends on γ, x, s and, in addition, to promised utility ϕ. We obtain the following equation, which is identical to (5.2) except that $\inf_\lambda \sup_a$ is replaced by $\sup_a \inf_\lambda$.

$$
\begin{aligned}
&W(\gamma, x, s, \phi) \\
&= \inf_{\mu \in \mathbb{R}_+^I} \sup_{a \in \tilde{\mathcal{A}}(x,s)} \inf_{\lambda \in \mathbb{R}_+^I} \left[\left(r(x, a, s) + \sum_{i=1}^I \left(\gamma^i g^i(x, a, s) + \lambda^i(g^i(x, a, s) - \bar{g}^i) + \mu^i(g^i(x, a, s) - \phi^i) \right) \right) \right. \\
&\quad \left. + \beta\mathbb{E}_s D(\gamma + \lambda + \mu, x', s') \right],
\end{aligned}
$$

where $x' = \zeta(x, a, s)$.

(5.11)

In the algorithm, we can change $\sup_a \inf_\lambda$ to $\inf_{\lambda_a} \sup_a$ which brings us back to the exact same setting.

6 Application: A Ramsey problem

We give a simple example of a problem where two-period ex-post lotteries are optimal. The two-period version of the problem (4.16) with S states $s = 1, \ldots, S$ in the second period

$$
\max_{(c_s, \ell_s, \tau_s)_{s=0}^S, q, b} u(c_0) + v(\ell_0) + \beta\sum_{s=1}^S p_s\left(u(c_s) + v(\ell_s)\right), \tag{6.1}
$$

$$
\text{s.t.} \ \ c_0 = b_{-1} + (1 - \tau_0)\ell_0 - qb, \ \ c_s = b + (1 - \tau_s)\ell_s, \ s = 1, \ldots, S; \tag{6.2}
$$

$$
(1 - \tau_s)u'(c_s) = -v'(\ell_s), \quad s = 0, \ldots, S; \tag{6.3}
$$

$$
-qu'(c_0) + \beta\sum_{s=1}^S p_s u'(c_s) = 0; \tag{6.4}
$$

$$
qb = b_{-1} - \tau_0\ell_0 + g_0, \ \ 0 = b - \tau_s\ell_s + g_s, \ s = 1, \cdots, S, \tag{6.5}
$$

where p_s, $s = 1, \ldots, S$ denotes the probability of state s and b_{-1} gives the initial condition.

For simplicity, we assume $u(c) = \log c$ and $v(\ell) = \sqrt{(1 - \ell)}$, $\beta = 1$. The following lemma shows how the constraints of the Ramsey problem can be simplified.

Lemma 6.1 *We define*

$$
h(\ell) := 1 - \frac{\ell}{2\sqrt{1 - \ell}}, \ f(\ell; g) := (\ell - g)h(\ell). \tag{6.6}
$$

For any $b > 0$, the following are equivalent:

22

1. $\{(c_s, \ell_s, \tau_s)_{s=0}^{S}, q, b\}$ *satisfies* (6.2) *to* (6.5);

2. *The equations*

$$b = f(\ell_s; g_s), \quad \forall s \in \{1, \cdots, S\}, \tag{6.7}$$

and

$$b_{-1} = f(\ell_0; g_0) + (\ell_0 - g_0)(\underbrace{\beta p h(\ell_1) + \beta(1-p)h(\ell_2)}_{w}), \tag{6.8}$$

have solutions with $b > 0$, $\ell_s > g_s$, $\forall s \in \{0, \cdots, S\}$.

The functions $f(\ell; g)$ can be interpreted as the net surplus of the government that sets a tax rate to ensure labor supply of ℓ if expenditure is g. We will discuss its shape in a concrete example below.

We first introduce the model with ex-post lotteries. It is assumed that the government can randomize over tax rates in each state $s \in \{1, \cdots, S\}$ in period 1. To simplify the exposition suppose that, given $b \geq 0$, the government can randomize between a high tax rate τ_s^H (corresponding to agent's labor supply ℓ_s^L), and a lower tax rate τ_s^L (corresponding to agent's labor supply ℓ_s^H). The lottery problem is then written

$$\max_{(\ell_s > g_s)_{s=0}^{S}, (\pi_s), \bar{b} \geq b \geq 0} u(\ell_0 - g_0) + v(\ell_0) + \beta \sum_{s=1}^{S} p_s \sum_{i \in \{L,H\}} \pi_{is}(u(\ell_s^i - g_s) + v(\ell_s^i)),$$

s.t. (6.7), and

$$b_{-1} = f(\ell_0; g_0) + (\ell_0 - g_0)(\beta \underbrace{\sum_{s=1}^{S} \sum_{i \in \{L,H\}} p_s \pi_{is} h(\ell_s^i)}_{w'}). \tag{6.9}$$

To illustrate why lotteries can be benfitial we consider a concrete example and assume $p_1 = 0.9$, as well as $0 = g_0 = g_L < g_H = 0.65$. Figure 1 shows how much revenue the government can generate in each state s and the associated utility in that state. Note first that the amount of debt is limited by its ability to pay back in state two. Debt levels above around 0.025 are not feasible. Second, for a given level of feasible debt, the government can collect the necessary revenue via a very high tax which implies low ℓ and low welfare and a relatively low tax which implies high ℓ and high welfare. The key to why the government might find
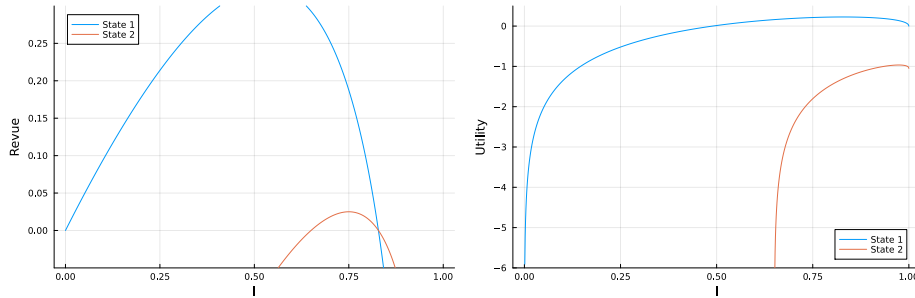

Figure 1: The left panel shows the the functions $f(\ell; g(s))$ for $s = 1, 2$. The right panel shows utility as a function of ℓ for $g = 0$.

it optimal to sometimes set very high taxes lies in the fact that it can only generate a large

revenue from selling bonds at $t = 0$ if the interest rate is sufficiently low. Since borrowing is constrained by the large government expenditure in state 2, $g_2 = 0.65$, the government might find it optimal to announce a very high tax for state 1, leading to very low consumption in state 1 and high demand for savings and very low interest rates in period zero. Formally, the revenue in period zero from debt is given by $(\ell_0 - g_0)\beta \sum_{s=1}^{S} \sum_{i \in \{L, H\}} p_s^i h(\ell_s^i)$ and the function $h(\ell)$ is equal to one at $\ell = 0$ and strictly decreasing in ℓ.

In this example, the government's ability to raise revenue from debt in period zero turns out to be extremely limited if it chooses the high-welfare low tax rates in state 1 in the future. Figure 2 shows a scatter plot of possible combinations of revenue at $t = 0$ and overall welfare. The left hand panel takes the tax rate in state 2 to be high, i.e. ℓ_2^L, the right hand panel takes it to be low. The figure shows that after a certain threshold (at about 0.3) the government can only raise the desired revenue from bond sales and first period taxes if it commits to a very high tax rate in state 1 in the second period. The resulting welfare drops dramatically. The differences between ℓ_2^H and ℓ_2^L of feasible welfare-revenue combinations can be seen to be minor, setting a low tax rate (resulting in ℓ_2^H) in the second period generally dominates slightly.

The right panel of Figure 2 also shows the resulting revenue-welfare combinations if the government randomizes (with probability 0.5) between ℓ_1^H and ℓ_1^L. For the case where it has to raise relatively high revenue, this clearly dominates the lottery-free solution. In particular, it can be seen in the figure (and verified numerically) that for b_- between 0.4 and 0.5 a lottery solution strictly dominates all lottery-free solutions as it generates the necessary revenue while leading to much higher welfare.

It is subject to further research how prevelant this problem is in realistically calibrated infinite-horizon versions of the model. The results in Citanna and Tvede (2024) indicate that it might be quite relevant. However, in the context of our simple example one might argue that in a stationary environment a forward looking government will never choose high debt levels.
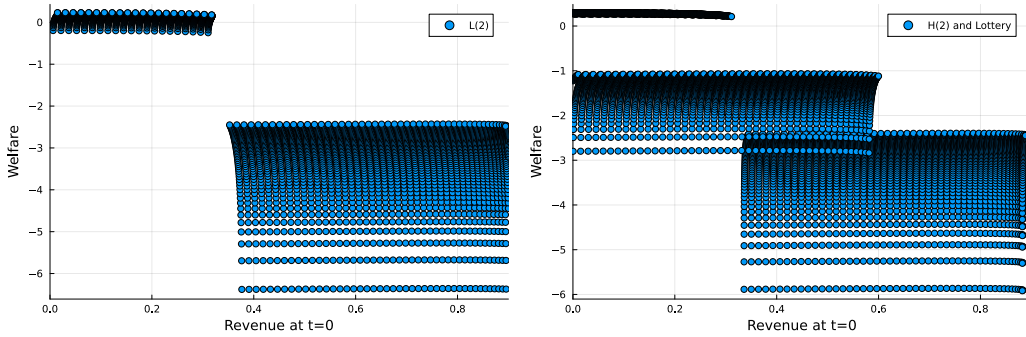

Figure 2: The left panel shows a scatter plot of feasible combinations of overall utility and period zero revenue for $\ell_2 = \ell^L$. The right panel shows a scatter plot of feasible combinations of overall utility and period zero revenue for $\ell_2 = \ell^H$, as well as for an equi-probable lottery between ℓ_1^H and ℓ_1^L

7 Conclusion

We show how the recursive multiplier approach to optimization problems with forward-looking contains can be used in a model with non-convexities. In general, neither the inf-sup nor the sup-inf formulation of the functional equation gives the correct solution for a model where

lotteries are not allowed. We argue that for many applications the sup-inf approach gives the value of an economically meaningful lottery solution.

It is subject to further research how to make our approach numerically viable in large scale models.

References

AIYAGARI, S. R., A. MARCET, T. J. SARGENT, AND J. SEPPÄLÄ (2002): "Optimal taxation without state-contingent debt," Journal of Political Economy, 110, 1220–1254.

ALIPRANTIS, C. D. AND K. C. BORDER (2006): Infinite dimensional analysis: a hitchhiker's guide, Springer.

ARNOTT, R. AND J. STIGLITZ (1988): "Randomization with Asymmetric Information," RAND Journal of Economics, 19, 344–362.

ATTANASIO, O. AND J.-V. RIOS-RULL (2000): "Consumption smoothing in island economies: Can public insurance reduce welfare?" European economic review, 44, 1225–1258.

BILLINGSLEY, P. (2013): Convergence of probability measures, John Wiley & Sons.

BLOISE, G. AND P. SICONOLFI (2022): "A Negishi Approach to Recursive Contracts," Econometrica, 90, 2821–2855.

BREZIS, H. AND H. BRÉZIS (2011): Functional analysis, Sobolev spaces and partial differential equations, vol. 2, Springer.

CITANNA, A. AND M. TVEDE (2024): "Do taxspots matter?: A study of optimal tax uncertainty," Journal of Economic Theory, 220, 105884.

COLE, H. AND F. KUBLER (2012): "Recursive contracts, lotteries and weakly concave pareto sets," Review of Economic Dynamics, 15, 479–500.

COOLEY, T., R. MARIMON, AND V. QUADRINI (2004): "Aggregate consequences of limited contract enforceability," Journal of political Economy, 112, 817–847.

DECHERT, W. D. (1982): "Lagrange multipliers in infinite horizon discrete time optimal control models," Journal of Mathematical Economics, 9, 285–302.

GAUTHIER, S. AND G. LAROQUE (2014): "On the value of randomization," Journal of Economic Theory, 151, 493–507.

KEHOE, P. J. AND F. PERRI (2002): "International business cycles with endogenous incomplete markets," Econometrica, 70, 907–928.

MARCET, A. AND R. MARIMON (2019): "Recursive contracts," Econometrica, 87, 1589–1631.

MARIMON, R. AND J. WERNER (2021): "The envelope theorem, Euler and Bellman equations, without differentiability," Journal of Economic Theory, 196, 105309.

MYERSON, R. B. (1982): "Optimal coordination mechanisms in generalized principal–agent problems," Journal of mathematical economics, 10, 67–81.

NEDIC, A. AND D. BERTSEKAS (2001): "Convergence rate of incremental subgradient algorithms," Applied Optimization, 54, 223–264.

PAVONI, N., C. SLEET, AND M. MESSNER (2018a): "The dual approach to recursive optimization: theory and examples," Econometrica, 86, 133–172.

——— (2018b): "Online Appendix The Dual Approach to Recursive Optimization: Theory and Examples," .

PAVONI, N. AND G. L. VIOLANTE (2007): "Optimal welfare-to-work programs," The Review of Economic Studies, 74, 283–318.

PRESCOTT, E. C. AND R. M. TOWNSEND (1984): "Pareto optima and competitive equilibria with adverse selection and moral hazard," Econometrica: journal of the econometric society, 21–45.

PRESCOTT, E. S. (1999): "A primer on moral-hazard models," FRB Richmond Economic Quarterly, 85, 47–77.

ROCKAFELLAR, R. T. (1974): Conjugate duality and optimization, SIAM.

——— (2015): "Convex analysis:(pms-28)," .

SHEN, C., F. KÜBLER, Y. YANG, AND Z. ZHOU (2025): "A Lagrangian Approach to Optimal Lotteries in Non-Convex Economies," arXiv preprint arXiv:2504.15997.

SPEAR, S. E. AND S. SRIVASTAVA (1987): "On repeated moral hazard with discounting," The Review of Economic Studies, 54, 599–617.

STIGLITZ, J. E. (1982): "Self-selection and Pareto efficient taxation," Journal of public economics, 17, 213–240.

STOKEY, N. L. AND R. E. LUCAS JR (1989): Recursive methods in economic dynamics, Harvard University Press.

URRUTY, J.-B. H. AND C. LEMARÉCHAL (1993): Convex analysis and minimization algorithms II, Springer-Verlag.

WEISS, L. (1976): "The desirability of cheating incentives and randomnness in the optimal income tax," Journal of Political Economy, 84, 1343–1352.

YOSIDA, K. AND E. HEWITT (1952): "Finitely additive measures," Transactions of the American Mathematical Society, 72, 46–66.

Appendix

A   Proofs

Proof of Lemma 3.1

It suffices to show that

$$\inf_{\boldsymbol{y}\in\Gamma}\sup_{x\in\tilde{X}} f(x,\boldsymbol{y}(x)) \leq \sup_{x\in\tilde{X}}\inf_{\boldsymbol{y}\in\Gamma} f(x,\boldsymbol{y}(x)). \tag{A.1}$$

For any $\epsilon > 0$ and $x \in \tilde{X}$, there exists $y_\epsilon(x) \in Y$, such that

$$f(x, y_\epsilon(x)) \leq \inf_{y\in Y} f(x,y) + \epsilon.$$

Therefore, taking $\boldsymbol{y} \in \Gamma$ s.t. $\boldsymbol{y}(x) = y_\epsilon(x)$ for any $x \in \tilde{X}$, we have

$$\sup_{x\in\tilde{X}} f(x, y_\epsilon(x)) \leq \sup_{x\in\tilde{X}}\inf_{y\in Y} f(x,y) + \epsilon$$
$$= \sup_{x\in\tilde{X}}\inf_{\boldsymbol{y}\in\Gamma} f(x,\boldsymbol{y}(x)) + \epsilon,$$

implying that

$$\inf_{\boldsymbol{y}\in\Gamma}\sup_{x\in\tilde{X}} f(x,\boldsymbol{y}(x)) \leq \sup_{x\in\tilde{X}}\inf_{\boldsymbol{y}\in\Gamma} f(x,\boldsymbol{y}(x)) + \epsilon.$$

Since ϵ is chosen arbitrarily, this implies (A.1). $\qquad\square$

Proof of Theorem 3.1

Using (3.3) and (3.4), we obtain the following.

$$D(\gamma, x_0, s_0) = \inf_{(\lambda_t^i(h^t))\in\Lambda}\sup_{(a_t)\in\tilde{\mathcal{A}}(x_0,s_0)} L((a_t),(\lambda_t^i(h^t));(\gamma^i),x_0,s_0)$$

$$= \inf_{(\lambda_t^i(h^t))\in\Lambda}\sup_{(a_t)\in\tilde{\mathcal{A}}(x_0,s_0)} \mathbb{E}_{s_0}\sum_{t=0}^{\infty}\beta^t\left(r(x_t,a_t,s_t)+\sum_{i=1}^{I}\gamma^i g^i(x_t,a_t,s_t)\right.$$

$$\left. + \sum_{i=1}^{I}\lambda_t^i(s_0,a_0,\cdots,s_{t-1},a_{t-1},s_t)\left(\sum_{n=0}^{\infty}\beta^n g^i(x_{t+n},a_{t+n},s_{t+n})-\bar{g}^i\right)\right)$$

$$= \inf_{(\lambda_0^i)}\inf_{(\lambda_t^i(h^t))_{t\geq1}}\sup_{a_0}\sup_{(a_t)_{t\geq1}}\underbrace{\left(r(a_0,x_0,s_0)+\sum_{i=1}^{I}\left(\gamma^i g^i(a_0,x_0,s_0)+\lambda_0^i(g^i(a_0,x_0,s_0)-\bar{g}^i)\right)\right)}_{h_1((\lambda_0^i),a_0)}$$

$$+ \underbrace{\mathbb{E}_{s_0}\sum_{t=1}^{\infty}\beta^t\left(r(a_t,x_t,s_t)+\sum_{i=1}^{I}\left((\gamma^i+\lambda_0^i)g^i(a_t,x_t,s_t)+\lambda_t^i(s^t,a^{t-1})\left(\sum_{n=0}^{\infty}\beta^n g^i(a_{t+n},x_{t+n},s_{t+n})-\bar{g}^i\right)\right)\right)}_{h_2((\lambda_0^i),a_0,(a_t)_{t\geq1},(\lambda_t^i(s^t,a^{t-1}))_{t\geq1})},$$

$$\tag{A.2}$$

where the third equation separates the terms concerning $t = 0$ and $t \geq 1$. [11]. We define $X = \mathcal{A}$, $\tilde{X} = \mathcal{A}(x_0, s_0)$ and $Y = \Lambda$. Since $(\lambda_t^i(h^t))_{t \geq 1} = (\lambda_t^i(s_0, a_0, h^{t-1}))_{t \geq 1}$, we can view it as a map from X to Y for each $s_0 \in \mathcal{S}$, and $\sup_{(a(s^t))_{t \geq 1}} h_1((\lambda_0^i), a_0) + h_2((\lambda_0^i), a_0, (a(s^t))_{t \geq 1}, (\lambda_t^i(s^t, a^{t-1}))_{t \geq 1})$ can be regarded as a functional $f(a_0, (\lambda_t^i(s_0, a_0, h^{t-1}))_{t \geq 1})$. By Lemma 3.1, we can exchange the operator $\inf_{(\lambda_t(h^t))_{t \geq 1}}$ and $\sup_{a_0}$ to obtain

$$
D(\gamma, x_0, s_0) = \inf_{(\lambda_0^i)} \inf_{(\lambda_t^i(h^t))_{t \geq 1}} \sup_{a_0} \sup_{(a(s^t))_{t \geq 1}} h_1((\lambda_0^i), a_0) + h_2((\lambda_0^i), a_0, (a_t)_{t \geq 1}, (\lambda_t^i(s^t, a^{t-1}))_{t \geq 1})
$$

$$
= \inf_{(\lambda_0^i)} \sup_{a_0} \left[h_1((\lambda_0^i), a_0) + \inf_{(\lambda_t^i(h^t))_{t \geq 1}} \sup_{(a_t)_{t \geq 1}} h_2((\lambda_0^i), a_0, (a_t)_{t \geq 1}, (\lambda_t^i(s^t, a^{t-1}))_{t \geq 1}) \right].
$$
(A.3)

For any possible realization of the shock at $t = 1$, $\bar{s}_1 \in \mathcal{S}_1$, we denote $a(s^t; \bar{s}_0, \bar{s}_1)$ the action at s^t if in history s_t, $s_0 = \bar{s}_0$, $s_1 = \bar{s}_1$, similarly let $\lambda_t^i(h^t; \bar{s}_0, \bar{a}_0, \bar{s}_1)$ denotes the multiplier at history h^t after specific realizations of $\bar{s}_0, \bar{a}_0$ and $\bar{s}_1$. For each $\bar{s}_1 \in \mathcal{S}$, define

$$
h_3((a(s^t|(s_0, \bar{s}_1)))_{t \geq 1}, (\lambda_t^i(h^t|(s_0, a_0, \bar{s}_1)))_{t \geq 1}; s_0, \bar{s}_1) =
$$

$$
\mathbb{E}_{\bar{s}_1} \sum_{t=1}^{\infty} \beta^{t-1} \left(r(a(s^t; s_0, \bar{s}_1), x_t, s_t) + \sum_{i=1}^{I} \left((\gamma^i + \lambda_0^i) g^i(a(s^t; s_0, \bar{s}_1), x_t, s_t) + \lambda_t^i(s^{t+1}; s_0, a_0, \bar{s}_1) \left(\sum_{n=0}^{\infty} \beta^n g^i(a(s^{t+n}; s_0, \bar{s}_1), x_{1+t}, s_{t+n}) - \bar{g}^i \right) \right) \right)
$$

We then have

$$
\inf_{(\lambda_t^i(h^t))_{t \geq 1}} \sup_{(a_t)_{t \geq 1}} h_2((\lambda_0^i), a_0, (a_t)_{t \geq 1}, (\lambda_t^i(s^t, a^{t-1}))_{t \geq 1})
$$

$$
= \inf_{(\lambda_t^i(h^t))_{t \geq 1}} \sup_{(a_t)_{t \geq 1}} \mathbb{E}_{s_0} \sum_{t=1}^{\infty} \beta^t \left(r(a_t, x_t, s_t) + \sum_{i=1}^{I} (\gamma^i + \lambda_0^i) g^i(a_t, x_t, s_t) + \lambda_t^i(s^t, a^{t-1}) \left(\sum_{n=0}^{\infty} \beta^n g^i(a_{t+n}, x_{t+n}, s_{t+n}) - \bar{g}^i \right) \right)
$$

$$
= \inf_{(\lambda_t^i(h^t|(s_0, a_0, \bar{s}_1)))_{t \geq 1, \bar{s}_1 \in \mathcal{S}}} \sup_{(a_t(s^t|(s_0, \bar{s}_1)))_{t \geq 1, \bar{s}_1 \in \mathcal{S}}} \beta \sum_{s_1 \in \mathcal{S}} \pi(\bar{s}_1|s_0) h_3((a_t(s^t|(s_0, \bar{s}_1)))_{t \geq 1}, (\lambda_t^i(h^t|(s_0, a_0, \bar{s}_1)))_{t \geq 1}; s_0, \bar{s}_1)
$$

$$
= \beta \sum_{\bar{s}_1 \in \mathcal{S}} \pi(\bar{s}_1|s_0) \inf_{(\lambda_t^i(h^t|(s_0, a_0, \bar{s}_1)))_{t \geq 1}} \sup_{(a(s^t|(s_0, \bar{s}_1)))_{t \geq 1}} h_3((a_t(s^t|(s_0, \bar{s}_1)))_{t \geq 1}, (\lambda_t^i(h^t|(s_0, a_0, \bar{s}_1)))_{t \geq 1}; s_0, \bar{s}_1)
$$

$$
= \beta \sum_{s_1 \in \mathcal{S}} \pi(s_1|s_0) D(\gamma + \lambda_0, x_1, s_1)
$$

$$
= \beta \mathbb{E}_{s_0} D(\gamma + \lambda_0, x_1, s_1),
$$
(A.4)

where the first equation separates the terms concerning different $s_1 \in \mathcal{S}$, and the second equation uses the definition of the dual value function D[12]. The recursive formulation (3.5) can then be obtained by combining (A.3) and (A.4). □

---

[11] In the expression $\sup_{a_0} \sup_{(a_t)_{t \geq 1}}$, we omit the feasible sets for a_0 and $(a_t)_{t \geq 1}$. Specifically:

- $a_0 \in \tilde{\mathcal{A}}(x_0, s_0)$,
- $(a_t(s_0, s_1, s^{t-2}))_{t \geq 1} \in \tilde{\mathcal{A}}^{\infty}(x_1, s_1)$, where $x_1 = \zeta(x_0, a_0(s_0), s_0)$, for all $s_0, s_1 \in \mathcal{S}$.

[12] Note that D is finite according to the fifth assumption in Assumption 3.1, which is important to ensure the finiteness of the sum of several dual value functions

Proof of Theorem 3.2

Following the same procedure as in the proof for Theorem 3.1, we have

$$\tilde{D}(\gamma, x, s) = \inf_{\lambda \in \mathcal{A} \to \mathbb{R}_+^I} \sup_{a \in \mathcal{A}} \left[\left(r(x, a, s) + \sum_{i=1}^{I} \left(\gamma^i g^i(x, a, s) + \lambda^i(a)(g^i(x, a, s) - \bar{g}^i) \right) \right) + \beta \mathbb{E}_s \tilde{D}(\gamma + \lambda(a), x', s') \right],$$

where $x' = \zeta(x, a, s),$ and $p(x, a, s) \geq 0.$

$$(A.5)$$

It then follows from Lemma 3.1 that

$$\tilde{D}(\gamma, x, s) = \inf_{\lambda \in \mathcal{A} \to \mathbb{R}_+^I} \sup_{a \in \mathcal{A}} \left[\left(r(x, a, s) + \sum_{i=1}^{I} \left(\gamma^i g^i(x, a, s) + \lambda^i(a)(g^i(x, a, s) - \bar{g}^i) \right) \right) + \beta \mathbb{E}_s \tilde{D}(\gamma + \lambda(a), x', s') \right]$$

$$= \sup_{a \in \mathcal{A}} \inf_{\lambda \in \mathbb{R}_+^I} \left[\left(r(x, a, s) + \sum_{i=1}^{I} \left(\gamma^i g^i(x, a, s) + \lambda^i(g^i(x, a, s) - \bar{g}^i) \right) \right) + \beta \mathbb{E}_s \tilde{D}(\gamma + \lambda, x', s') \right],$$

where $x' = \zeta(x, a, s),$ and $p(x, a, s) \geq 0.$

$$\square$$

Proof of Lemma 3.3

- **Show that $\mathcal{B}$ preserves convexity.** See the proof of Lemma 2 in Pavoni et al. (2018a).

- **Show that $\mathcal{B}$ preserves L-Lipschitz continuity.** Assume that F is L-Lipschitz. Given $\gamma_1, \gamma_2 \in \mathbb{R}_+^I$, we have

$$\mathcal{B}(F)(x, \gamma_1, s) = \inf_{\lambda \in \mathbb{R}_+^I} \sup_{a \in \tilde{\mathcal{A}}(x, s)} \left[\left(r(a, x, s) + \sum_{i=1}^{I} \left(\gamma_1^i g^i(a, x, s) + \lambda^i(g^i(a, x, s) - \bar{g}^i) \right) \right) + \beta \mathbb{E}_s F(\gamma_1 + \lambda, x', s') \right]$$

$$= \inf_{\lambda \in \mathbb{R}_+^I} \sup_{a \in \tilde{\mathcal{A}}(x, s)} \left[\left(r(a, x, s) + \sum_{i=1}^{I} \left(\gamma_2^i g^i(a, x, s) + \lambda^i(g^i(a, x, s) - \bar{g}^i) \right) \right) + \beta \mathbb{E}_s F(\gamma_2 + \lambda, x', s') \right.$$

$$\left. + \sum_{i=1}^{I} (\gamma_1^i - \gamma_2^i) g^i(a, x, s) + \beta \mathbb{E}_s (F(\gamma_1 + \lambda, x', s') - F(\gamma_2 + \lambda, x', s')) \right]$$

Since

$$\sum_{i=1}^{I} (\gamma_1^i - \gamma_2^i) g^i(a, x, s) + \beta \mathbb{E}_s (F(\gamma_1 + \lambda, x', s') - F(\gamma_2 + \lambda, x', s'))$$

$$\leq \|\gamma_1 - \gamma_2\|_1 \max_i \|g^i\|_\infty + \beta L \|\gamma_1 - \gamma_2\|_1 \leq L \|\gamma_1 - \gamma_2\|_1,$$

29

we have

$$
\begin{aligned}
\mathcal{B}(F)(x, \gamma_1, s) &= \inf_{\lambda \in \mathbb{R}_+^I} \sup_{a \in \tilde{\mathcal{A}}(x,s)} \left[\left(r(a, x, s) + \sum_{i=1}^I \left(\gamma_2^i g^i(a, x, s) + \lambda^i (g^i(a, x, s) - \bar{g}^i) \right) \right) + \beta \mathbb{E}_s F(\gamma_2 + \lambda, x', s') \right. \\
&\quad \left. + \sum_{i=1}^I (\gamma_1^i - \gamma_2^i) g^i(a, x, s) + \beta \mathbb{E}_s (F(\gamma_1 + \lambda, x', s') - F(\gamma_2 + \lambda, x', s')) \right] \\
&\leq \inf_{\lambda \in \mathbb{R}_+^I} \sup_{a \in \tilde{\mathcal{A}}(x,s)} \left[\left(r(a, x, s) + \sum_{i=1}^I \left(\gamma_2^i g^i(a, x, s) + \lambda^i (g^i(a, x, s) - \bar{g}^i) \right) \right) + \beta \mathbb{E}_s F(\gamma_2 + \lambda, x', s') \right. \\
&\quad \left. + L\|\gamma_1 - \gamma_2\|_1 \right] = \mathcal{B}(F)(x, \gamma_2, s) + L\|\gamma_1 - \gamma_2\|_1
\end{aligned}
$$

Hence $\mathcal{B}$ preserves L-Lipschitz continuity.

- **Show that $\mathcal{B}$ preserves the property** (3.10). Assume that F satisfies (3.10). Given $x \in \mathcal{X}$, $\gamma \in \mathbb{R}_+^I$, $s \in \mathcal{S}$. Assume that $(a(s^t))_{t \geq 0}$ is a feasible control to problem (3.1) with $x_0 = x$, $s_0 = s$, satisfying

$$
\begin{cases}
v^0 = \mathbb{E}_{s_0} \sum_{t=0}^\infty \beta^t r(x_t, a(s^t), s_t); \\
v^i = \mathbb{E}_{s_0} \sum_{t=0}^\infty \beta^t g^i(x_t, a(s^t), s_t).
\end{cases}
$$

We denote $x' = \zeta(a(s_0), x_0, s_0)$, it is straightforward to verify that $a(s^t|(s, s'))$ is a feasible control to problem (3.1) with $x_0 = x'$, $s_0 = s'$. Therefore, for any $\lambda \in \mathbb{R}_+^I$, we have

$$
\sup_{a \in \mathcal{A}} \left[\left(r(a, x, s) + \sum_{i=1}^I \left(\gamma^i g^i(a, x, s) + \lambda^i (g^i(a, x, s) - \bar{g}^i) \right) \right) + \beta \mathbb{E}_s F(\gamma + \lambda, x', s') \right]
$$

$$
\begin{aligned}
(F \text{ satisfies } (3.10)) &\geq r(a_0(s_0), x_0, s_0) + \sum_{i=1}^I \left(\gamma^i g^i(a_0(s_0), x_0, s_0) + \lambda^i (g^i(a_0(s_0), x_0, s_0) - \bar{g}^i) \right) \\
&\quad + \beta \mathbb{E}_{s_0} \sum_{t=1}^\infty \beta^{t-1} r(x_t, a_t(s^t), s_t) + \sum_{i=1}^I (\gamma^i + \lambda^i) \beta^{t-1} g^i(x_t, a_t(s^t), s_t) \\
&= v^0 + \sum_{i=1}^I \left(\gamma^i v^i + \lambda^i (v^i - \bar{g}^i) \right) \\
&\geq v^0 + \sum_{i=1}^I \gamma^i v^i.
\end{aligned}
$$

Hence $\mathcal{B}(F)(\gamma, x, s) \geq v^0 + \sum_{i=1}^I \gamma^i v^i$. Since $x \in \mathcal{X}$, $\gamma \in \mathbb{R}_+^I$, $s \in \mathcal{S}$, and the feasible control $(a_t(s^t))_{t \geq 0}$ with $x_0 = x$, $s_0 = s$ are chosen arbitrarily, it is then straightforward to conclude that $\mathcal{B}(F)$ satisfies (3.10).

- **Show that $\mathcal{B}$ preserves the property** (3.12). By the definition of $\mathcal{B}$, we have

$$
\begin{aligned}
\mathcal{B}(F)(x,\gamma,s) &= \inf_{\lambda \in \mathbb{R}_+^I} \sup_{a \in \tilde{\mathcal{A}}(x,s)} \left[\left(r(a,x,s) + \sum_{i=1}^I \left(\gamma^i g^i(a,x,s) + \lambda^i(g^i(a,x,s) - \bar{g}^i) \right) \right) + \beta \mathbb{E}_s F(\gamma+\lambda, x', s') \right] \\
&\leq \sup_{a \in \tilde{\mathcal{A}}(x,s)} \left[\left(r(a,x,s) + \sum_{i=1}^I \gamma^i g^i(a,x,s) \right) + \beta \mathbb{E}_s F(\gamma, x', s') \right] \\
&\leq \|r\|_\infty + \sum_{i=1}^I \gamma^i \|g_i\|_\infty + \beta \left(1 + \sum_{i=1}^I \gamma^i \right) L \\
&\leq \left(1 + \sum_{i=1}^I \gamma^i \right) L,
\end{aligned}
$$

  where the last step uses the identity $(1-\beta)L = \|r\|_\infty + \sum_i \|g_i\|_\infty$ together with non-negativity.

- **Show that $\mathcal{B}$ preserves the finiteness of the norm $\|\cdot\|_{\mathcal{M}}$.** This is a direct corollary of the fact that $\mathcal{B}(F)$ preserves the properties (3.10) and (3.12). $\qquad \square$

Proof of Theorem 3.3

- **Step 1. Show that there exists a pointwise limit of $\mathcal{B}^{(n)}(F_0)$ in $\mathcal{N}$.** According to Lemma 3.3, since $F_0 \in \mathcal{N}$, we know that

$$
\mathcal{B}(F_0) \in \mathcal{N},
$$

  implying that $\mathcal{B}(F_0) \leq F_0$. By induction and monotonicity we have

$$
\mathcal{B}^{(n)}(F_0) \leq \mathcal{B}^{(n-1)}(F_0), \ \forall n \in \mathbb{N}_+.
$$

  Therefore, for any $\gamma \in \mathbb{R}_+^I$, $x \in \mathcal{X}$, $s \in \mathcal{S}$, the sequence

$$
\{\mathcal{B}^{(n)}(F_0)(\gamma, x, s)\}_{n \in \mathbb{N}_+}
$$

  is a decreasing, bounded sequence, and hence has a limit $F^*(\gamma, x, s)$. It is straightforward to verify that $F^* \in \mathcal{N}$.

- **Step 2. Show that $\mathcal{B}^{(n)}(F_0)$ converges to F^* in norm $\|\cdot\|_{\mathcal{M}}$.** For any $x \in \mathcal{X}$, $s \in \mathcal{S}$, we know that

$$
\mathcal{B}^{(n)}(F_0)(\gamma, x, s)
$$

  is L-Lipschitz and uniformly bounded in any compact ball $B(k)$, according to Arzela-Ascoli theorem, there exists a subsequence

$$
\{\mathcal{B}^{(t_n)}(F_0)(\gamma, x, s)\} \subset \{\mathcal{B}^{(n)}(F_0)(\gamma, x, s)\},
$$

s.t.
$$\mathcal{B}^{(t_n)}(F_0)(\cdot, x, s) \rightrightarrows \tilde{F}_k(\cdot, x, s), \text{ as } n \to \infty \text{ in } B(k),$$

for some $\tilde{F}_k(\cdot, x, s)$, where $\rightrightarrows$ denotes uniformly convergence. According to step 1, we know that
$$\lim_{n \to \infty} \mathcal{B}^{(t_n)}(F_0)(\gamma, x, s) = F^*(\gamma, x, s), \forall \gamma \in B(k).$$

Therefore, we have $\tilde{F}_k(\gamma, x, s) = F^*(\gamma, x, s)$, $\forall \gamma \in B(k)$, and hence

$$\mathcal{B}^{(t_n)}(F_0)(\cdot, x, s) \rightrightarrows F^*(\cdot, x, s), \text{ as } n \to \infty \text{ in } B(k).$$

According to the monotonicity, the whole sequence uniformly converges, i.e.

$$\mathcal{B}^{(n)}(F_0)(\cdot, x, s) \rightrightarrows F^*(\cdot, x, s), \text{ as } n \to \infty \text{ in } B(k),$$

yielding that

$$\|\mathcal{B}^{(n)}(F_0)(\cdot, x_j, s_i) - F^*(\cdot, x_j, s_i)\|_{L^\infty(B(k))} \to 0, \forall s_i \in \mathcal{S}, x_j \in \mathcal{X}, k \in \mathbb{N}_+. \tag{A.6}$$

Since functions in $\mathcal{N}$ is bounded by affine functions, there exists $C > 0$, s.t.

$$\|\mathcal{B}^{(n)}(F_0)(\cdot, x_j, s_i) - F^*(\cdot, x_j, s_i)\|_{L^\infty(B(k))} \le C(k+1), \forall s_i \in \mathcal{S}, x_j \in \mathcal{X}, k \in \mathbb{N}_+. \tag{A.7}$$

Therefore, applying (A.7) we have

$$\|\mathcal{B}^{(n)}(F_0) - F^*\|_{\mathcal{M}}$$
$$= \sum_{i=1}^{|\mathcal{S}|} \frac{1}{2^i} \left[\sum_{j=1}^{\infty} \frac{1}{2^j} \left(\sum_{k=1}^{\infty} \frac{1}{2^k} \|\mathcal{B}^{(n)}(F_0)(\cdot, x_j, s_i) - F^*(\cdot, x_j, s_i)\|_{L^\infty(B(k))} \right) \right]$$
$$\le \sum_{i=1}^{|\mathcal{S}|} \frac{1}{2^i} \left\{ \sum_{j=1}^{J} \frac{1}{2^j} \left[\left(\sum_{k=1}^{K} \frac{1}{2^k} \|\mathcal{B}^{(n)}(F_0)(\cdot, x_j, s_i) - F^*(\cdot, x_j, s_i)\|_{L^\infty(B(k))} \right) + \underbrace{\left(\sum_{k>K} \frac{(k+1)C}{2^k} \right)}_{I} \right] \right\}$$
$$+ \underbrace{\sum_{i=1}^{|\mathcal{S}|} \frac{1}{2^i} \sum_{j>J} \frac{1}{2^j} \sum_{k=1}^{\infty} \frac{(k+1)C}{2^k}}_{II}$$

For any $\epsilon > 0$, there exists $J > 0$ and $K > 0$, s.t. $I < \epsilon/3$, $II < \epsilon/3$. Furthermore, according to (A.6), there exists $N \in \mathbb{N}_+$, s.t. when $n > N$, we have

$$\sum_{i=1}^{|\mathcal{S}|} \frac{1}{2^i} \sum_{j=1}^{J} \frac{1}{2^j} \left[\left(\sum_{k=1}^{K} \frac{1}{2^k} \|\mathcal{B}^{(n)}(F_0)(\cdot, x_j, s_i) - F^*(\cdot, x_j, s_i)\|_{L^\infty(B(k))} \right) \right] < \frac{\epsilon}{3}.$$

32

Therefore, when $n > N$, we have

$$\|\mathcal{B}^{(n)}(F_0) - F^*\|_{\mathcal{M}}$$

$$\leq \sum_{i=1}^{|\mathcal{S}|} \frac{1}{2^i} \left\{ \sum_{j=1}^{J} \frac{1}{2^j} \left[\left(\sum_{k=1}^{K} \frac{1}{2^k} \|\mathcal{B}^{(n)}(F_0)(\cdot, x_j, s_i) - F^*(\cdot, x_j, s_i)\|_{L^\infty(B(k))} \right) + \frac{\epsilon}{3} \right] \right\} + \frac{\epsilon}{3}$$

$$\leq \sum_{i=1}^{|\mathcal{S}|} \frac{1}{2^i} \sum_{j=1}^{J} \frac{1}{2^j} \left[\left(\sum_{k=1}^{K} \frac{1}{2^k} \|\mathcal{B}^{(n)}(F_0)(\cdot, x_j, s_i) - F^*(\cdot, x_j, s_i)\|_{L^\infty(B(k))} \right) \right] + \frac{2\epsilon}{3} < \epsilon$$

Therefore $\|\mathcal{B}^{(n)}(F_0) - F^*\|_{\mathcal{M}} \to 0$ by definition.

Step 3. Show that $\mathcal{B}(F^*) = F^*$. According to monotonicity, we have

$$\mathcal{B}(F^*) \leq \mathcal{B}^{(n)}(F_0), \ \forall n \in \mathbb{N}_+.$$

For any $\gamma \in \mathbb{R}_+^I$, $x \in \mathcal{X}$, $s \in \mathcal{S}$, we take $n \to \infty$ and have

$$\mathcal{B}(F^*) \leq F^*.$$

If there exists $\ell > 0$ s.t. $\mathcal{B}(F^*)(\gamma, x, s) = F^*(\gamma, x, s) - \ell$ for some $x \in \mathcal{X}$, $s \in \mathcal{S}$, $\gamma \in \mathbb{R}_+^I$.
According to the definition of $\mathcal{B}$, we know that there exists $\lambda \in \mathbb{R}_+^I$, s.t.

$$\sup_{a \in \tilde{A}(x,s)} \left[\left(r(a, x, s) + \sum_{i=1}^{I} \gamma^i g^i(a, x, s) + \lambda^i(g^i(a, x, s) - \bar{g}^i) \right) + \beta \mathbb{E}_s F^*(\gamma + \lambda, x', s') \right] < F^*(\gamma, x, s) - \frac{\ell}{2}.$$

$$(A.8)$$

Since $\mathcal{B}^{(n)}(F_0) \to F^*$, there exists $N \in \mathbb{N}_+$, s.t.

$$\mathcal{B}^{(N)}(F_0)(\gamma + \lambda, x', s') - F^*(\gamma + \lambda, x', s') \leq \frac{\ell}{4}, \ \forall x' \in \{\zeta(x, a, s) | a \in \tilde{A}(x, s)\}, \ s' \in \mathcal{S}.$$

Therefore,

$$\mathcal{B}^{(N+1)}(F_0)(\gamma, x, s)$$

$$\leq \sup_{a \in \tilde{A}(x,s)} \left[\left(r(a, x, s) + \sum_{i=1}^{I} \gamma^i g^i(a, x, s) + \lambda^i(g^i(a, x, s) - \bar{g}^i) \right) + \beta \mathcal{B}^{(N)}(F_0)(\gamma + \lambda, x', s') \right]$$

$$< \sup_{a \in \tilde{A}(x,s)} \left[\left(r(a, x, s) + \sum_{i=1}^{I} \gamma^i g^i(a, x, s) + \lambda^i(g^i(a, x, s) - \bar{g}^i) \right) + \beta F^*(\gamma + \lambda, x', s') + \frac{\ell}{4} \right] < F^*(\gamma, x, s) - \frac{\ell}{4},$$

contradicting the fact that $\mathcal{B}^{(N+1)}(F_0)(\gamma, x, s) \geq F^*(\gamma, x, s)$. Hence $\mathcal{B}(F^*) \geq F^*$, and therefore $\mathcal{B}(F^*) = F^*$.

- **Step 4. Show that F^* is the largest fixed point in $\mathcal{N}$.** For any $F \in \mathcal{N}$, which is a fixed point of $\mathcal{B}$, we know that $F \leq F_0$. Therefore, according to the monotonicity of $\mathcal{B}$, we have

$$F = \mathcal{B}^{(n)}(F) \leq \mathcal{B}^{(n)}(F_0), \ \forall n \in \mathbb{N}_+.$$

33

We take $n \to \infty$ and have $F \leq F^*$. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Proof of Lemma 3.4

According to the definition of the dual value function $D(\gamma, x_0, s_0)$, for any $\epsilon > 0$, there exists $(\lambda_t^i(h^t)) \in \Lambda$, s.t.

$$\sup_{a_t(s^t) \in \tilde{\mathcal{A}}^\infty(x_0)} L((a_t), (\lambda_t^i); (\gamma^i), x_0, s_0) \leq D(\gamma, x_0, s_0) + \epsilon. \tag{A.9}$$

We define $M = \sum_{t=0}^\infty \sum_{h^t \in \mathcal{H}^t} \sum_{i=1}^I \beta^t \lambda_t^i(h^t) \pi^t(s^t|s_0)$.

- **Step 1.** We first show that,

$$\sup_{(a_t) \in \tilde{\mathcal{A}}^\infty(x_0)} L((a_t), (\lambda_t^i); (\gamma^i), x_0, s_0)$$

$$= \lim_{T \to \infty} \sup_{(a_t) \in \tilde{\mathcal{A}}^\infty(x_0)} \mathbb{E}_{s_0} \sum_{t=0}^T \beta^t \left(r(x_t, a_t, s_t) + \sum_{i=1}^I \gamma^i g^i(x_t, a_t, s_t) \right. \tag{A.10}$$

$$\left. + \lambda_t^i(s_0, a_0, \cdots, s_{t-1}, a_{t-1}, s_t) \left(\sum_{n=0}^{T-t} \beta^n g^i(x_{t+n}, a_{t+n}, s_{t+n}) - \bar{g}^i \right) \right).$$

Since r and (g^i) are bounded, for any $T \geq 0$ we have

$$\sup_{(a_t) \in \tilde{\mathcal{A}}^\infty(x_0)} L((a_t), (\lambda_t^i); (\gamma^i), x_0, s_0)$$

$$= \sup_{(a_t) \in \tilde{\mathcal{A}}^\infty(x_0)} \mathbb{E}_{s_0} \sum_{t=0}^\infty \beta^t \left(r(x_t, a_t, s_t) + \sum_{i=1}^I \gamma^i g^i(x_t, a_t, s_t) \right.$$

$$\left. + \lambda_t^i(s_0, a_0, \cdots, s_{t-1}, a_{t-1}, s_t) \left(\sum_{n=0}^\infty \beta^n g^i(x_{t+n}, a_{t+n}, s_{t+n}) - \bar{g}^i \right) \right)$$

$$= \sup_{(a_t) \in \tilde{\mathcal{A}}^\infty(x_0)} \mathbb{E}_{s_0} \underbrace{\sum_{t=0}^T \beta^t \left(r(x_t, a_t, s_t) + \sum_{i=1}^I \gamma^i g^i(x_t, a_t, s_t) \right.}_{I}$$

$$\underbrace{\left. + \lambda_t^i(s_0, a_0, \cdots, s_{t-1}, a_{t-1}, s_t) \left(\sum_{n=0}^{T-t} \beta^n g^i(x_{t+n}, a_{t+n}, s_{t+n}) - \bar{g}^i \right) \right)}_{I}$$

$$+ \underbrace{\sum_{t=T+1}^\infty \beta^t \left(r(x_t, a_t, s_t) + \sum_{i=1}^I \gamma^i g^i(x_t, a_t, s_t) \right.}_{II}$$

$$\underbrace{\left. + \lambda_t^i(s_0, a_0, \cdots, s_{t-1}, a_{t-1}, s_t) \left(\sum_{n=0}^\infty \beta^n g^i(x_{t+n}, a_{t+n}, s_{t+n}) - \bar{g}^i \right) \right)}_{II}$$

$$+ \underbrace{\sum_{t=0}^T \beta^t \lambda_t^i(s_0, a_0, \cdots, s_{t-1}, a_{t-1}, s_t) \left(\sum_{n=T-t+1}^\infty \beta^n g^i(x_{t+n}, a_{t+n}, s_{t+n}) \right)}_{III}.$$

To prove (A.10), it suffices to show that $II \to 0$ and $III \to 0$ uniformly for any $(a_t) \in \tilde{\mathcal{A}}^\infty(x_0)$ as $T \to \infty$. For II, according to the boundedness of r and g^i,

$$\mathbb{E}_{s_0} \sum_{t=T+1}^\infty \beta^t \left(r(x_t, a_t, s_t) + \sum_{i=1}^I \gamma^i g^i(x_t, a_t, s_t) \right) \to 0 \tag{A.11}$$

uniformly w.r.t. (a_t). Furthermore, since $\mathbb{E}_{s_0} \sum_{t=T+1}^\infty \sum_{h^t \in \mathcal{H}^t} \sum_{i=1}^I \beta^t \lambda_t^i(h^t) \to 0$, and that $|\sum_{n=0}^\infty \beta^n g^i(x_{t+n}, a_{t+n}, s_{t+n}) - \bar{g}^i|$ are bounded above by $\frac{\|g^i\|_\infty}{1-\beta} + \|\bar{g}^i\|_\infty$, we have

$$\mathbb{E}_{s_0} \sum_{t=T+1}^\infty \beta^t \sum_{i=1}^I \left(\lambda_t^i(s_0, a_0, \cdots, s_{t-1}, a_{t-1}, s_t) \left(\sum_{n=0}^\infty \beta^n g^i(x_{t+n}, a_{t+n}, s_{t+n}) - \bar{g}^i \right) \right) \to 0$$
$$\tag{A.12}$$

uniformly w.r.t. (a_t). Combining (A.11) and (A.12) we know that $II \to 0$ as $T \to \infty$ uniformly w.r.t. (a_t).

For III, we have

$$III = \mathbb{E}_{s_0} \underbrace{\sum_{t=0}^{[\frac{T}{2}]} \sum_{i=1}^{I} \beta^t \lambda_t^i(s_0, a_0, \cdots, s_{t-1}, a_{t-1}, s_t) \left(\sum_{n=T-t+1}^{\infty} \beta^n g^i(x_{t+n}, a_{t+n}, s_{t+n}) \right)}_{III(1)}$$

$$+ \underbrace{\sum_{t=[\frac{T}{2}]+1}^{T} \sum_{i=1}^{I} \beta^t \lambda_t^i(s_0, a_0(s_0), \cdots, s_{t-1}, a_{t-1}, s_t) \left(\sum_{n=T-t+1}^{\infty} \beta^n g^i(x_{t+n}, a_{t+n}, s_{t+n}) \right)}_{III(2)}.$$

Since $\mathbb{E}_{s_0} \sum_{t=0}^{[\frac{T}{2}]} \sum_{i=1}^{I} \beta^t \lambda_t^i(s_0, a_0, \cdots, s_{t-1}, a_{t-1}, s_t)$ is bounded above by M, and $\left(\sum_{n=T-t+1}^{\infty} \beta^n g^i(x_{t+n}, a_{t+n}, s_{t+n}) \right)$ is bounded above by $\frac{\beta^{T-[\frac{T}{2}]+1}}{1-\beta} \|g^i\|_\infty$, we know that $III(1) \to 0$ uniformly w.r.t. (a_t) when $T \to \infty$.

Since

$$\mathbb{E}_{s_0} \sum_{t=[\frac{T}{2}]+1}^{T} \sum_{i=1}^{I} \beta^t \lambda_t^i(s_0, a_0, \cdots, s_{t-1}, a_{t-1}, s_t)$$

$$\leq \mathbb{E}_{s_0} \sum_{t=[\frac{T}{2}]+1}^{\infty} \sum_{i=1}^{I} \beta^t \lambda_t^i(s_0, a_0, \cdots, s_{t-1}, a_{t-1}, s_t) \to 0,$$

uniformly w.r.t. (a_t), and $|\sum_{n=T-t+1}^{\infty} \beta^n g^i(x_{t+n}, a_{t+n}, s_{t+n})|$ are bounded above by $\frac{\|g_i\|_\infty}{1-\beta}$, we know that $III(2) \to 0$ uniformly w.r.t. (a_t) when $T \to \infty$. Hence $III \to 0$ uniformly w.r.t. (a_t) as $T \to \infty$, and therefore, (A.10) holds.

- **Step 2.** We then show that, given $x_0 \in \mathcal{X}$, $s_0 \in \mathcal{S}$, we have

$$\limsup_{T \to \infty} \sup_{(a_t(s^t)) \in \tilde{\mathcal{A}}^\infty(x_0)} \mathbb{E}_{s_0} \beta^{T+1} F(\gamma + \sum_{t=0}^{T} \lambda_t(s_0, a_0, \cdots, s_{t-1}, a_{t-1}, s_t), x_{T+1}, s_{T+1}) \leq 0. \tag{A.13}$$

According to (3.12), we know that for any $(a_t(s^t)) \in \tilde{\mathcal{A}}^\infty(x_0)$, we have

$$\mathbb{E}_{s_0} \beta^{T+1} F(\gamma + \sum_{t=0}^{T} \lambda_t(s_0, a_0, \cdots, s_{t-1}, a_{t-1}, s_t), x_t, s_t)$$

$$\leq \mathbb{E}_{s_0} \beta^{T+1} \left[1 + \sum_{i=1}^{I} \left(\gamma^i + \sum_{t=0}^{T} \lambda_t^i(s_0, a_0, \cdots, s_{t-1}, a_{t-1}, s_t) \right) \right] L$$

$$= \beta^{T+1} \left(1 + \sum_{i=1}^{I} \gamma^i \right) L + \beta^{T+1} L \mathbb{E}_{s_0} \sum_{t=0}^{T} \sum_{i=1}^{I} \lambda_t^i(s_0, a_0, \cdots, s_{t-1}, a_{t-1}, s_t)$$

$$\leq \underbrace{\beta^{T+1} \left(1 + \sum_{i=1}^{I} \gamma^i \right) L}_{IV} + \underbrace{\beta^{T+1} L \sum_{t=0}^{T} \sum_{h^t \in \mathcal{H}^t} \sum_{i=1}^{I} \pi(s^t|s_0) \lambda_t^i(h^t)}_{V}$$

36

Obviously $IV \to 0$ as $T \to \infty$, and

$$
\begin{aligned}
V = & L \sum_{t=0}^{T} \beta^{T+1-t} \sum_{h^t \in \mathcal{H}^t} \sum_{i=1}^{I} \left(\beta^t \pi(s^t|s_0) \lambda_t^i(h^t) \right) \\
\leq & L \sum_{t=0}^{[\frac{T}{2}]} \beta^{T+1-[\frac{T}{2}]} \sum_{h^t \in \mathcal{H}^t} \sum_{i=1}^{I} \left(\beta^t \pi(s^t|s_0) \lambda_t^i(h^t) \right) + L \sum_{t=[\frac{T}{2}]+1}^{\infty} \sum_{h^t \in \mathcal{H}^t} \sum_{i=1}^{I} \beta^t \pi(s^t|s_0) \lambda_t^i(h^t) \\
\leq & L \beta^{T+1-[\frac{T}{2}]} M + L \sum_{t=[\frac{T}{2}]+1}^{\infty} \sum_{h^t \in \mathcal{H}^t} \sum_{i=1}^{I} \beta^t \pi(s^t|s_0) \lambda_t^i(h^t) \to 0 (T \to \infty).
\end{aligned}
$$

Furthermore, observing that IV and V are independent on $(a_t(s^t))$, and we obtain (A.13).

- **Step 3.** According to (A.9) and (A.10), there exists $T_1 > 0$, s.t. when $T \geq T_1$,

$$
\begin{aligned}
D(\gamma, x_0, s_0) + 2\epsilon \geq & \sup_{(a_t(s^t)) \in \tilde{\mathcal{A}}^\infty(x_0)} \mathbb{E}_{s_0} \sum_{t=0}^{T} \beta^t \left(r(x_t, a_t, s_t) + \sum_{i=1}^{I} \gamma^i g^i(x_t, a_t, s_t) \right. \\
& \left. + \lambda_t^i(s_0, a_0, \cdots, s_{t-1}, a_{t-1}, s_t) \left(\sum_{n=0}^{T-t} \beta^n g^i(x_{t+n}, a_{t+n}, s_{t+n}) - \bar{g}^i \right) \right).
\end{aligned}
\tag{A.14}
$$

Combining the inequality (A.14) and (A.13), we know that there exists $T_2 > T$, s.t. when $T > \max\{T_1, T_2\}$,

$$
\begin{aligned}
D(\gamma, x_0, s_0) + 3\epsilon \geq & \sup_{(a_t) \in \tilde{\mathcal{A}}^\infty(x_0)} \mathbb{E}_{s_0} \sum_{t=0}^{T} \beta^t \left(r(x_t, a_t, s_t) + \sum_{i=1}^{I} \gamma^i g^i(x_t, a_t, s_t) \right. \\
& \left. + \lambda_t^i(s_0, a_0(s_0), \cdots, s_{t-1}, a_{t-1}(s^{t-1}), s_t) \left(\sum_{n=0}^{T-t} \beta^n g^i(x_{t+n}, a_{t+n}, s_{t+n}) - \bar{g}^i \right) \right) \\
& + \sup_{(a_t) \in \tilde{\mathcal{A}}^\infty(x_0)} \mathbb{E}_{s_0} \beta^{T+1} F(\gamma + \sum_{t=0}^{T} \lambda_t(s_0, a_0, \cdots, s_{t-1}, a_{t-1}, s_t), x_{T+1}, s_{T+1}) \\
\geq & \sup_{(a_t) \in \tilde{\mathcal{A}}^\infty(x_0)} \left[\mathbb{E}_{s_0} \sum_{t=0}^{T} \beta^t \left(r(x_t, a_t, s_t) + \sum_{i=1}^{I} \gamma^i g^i(x_t, a_t, s_t) \right. \right. \\
& \left. + \lambda_t^i(s_0, a_0, \cdots, s_{t-1}, a_{t-1}, s_t) \left(\sum_{n=0}^{T-t} \beta^n g^i(x_{t+n}, a_{t+n}, s_{t+n}) - \bar{g}^i \right) \right) \\
& \left. + \mathbb{E}_{s_0} \beta^{T+1} F(\gamma + \sum_{t=0}^{T} \lambda_t(s_0, a_0, \cdots, s_{t-1}, a_{t-1}, s_t), x_{T+1}, s_{T+1}) \right].
\end{aligned}
\tag{A.15}
$$

37

We rearrange the right hand side of (A.15) and have

$$
\sup_{(a_t)\in\tilde{\mathcal{A}}^\infty(x_0)}\left[\mathbb{E}_{s_0}\sum_{t=0}^{T}\beta^t\left(r(x_t,a_t,s_t)+\sum_{i=1}^{I}\gamma^i g^i(x_t,a_t,s_t)\right.\right.
$$
$$
\left.+\lambda_t^i(s_0,a_0,\cdots,s_{t-1},a_{t-1},s_t)\left(\sum_{n=0}^{T-t}\beta^n g^i(x_{t+n},a_{t+n},s_{t+n})-\bar{g}^i\right)\right)
$$
$$
\left.+\mathbb{E}_{s_0}\beta^{T+1}F(\gamma+\sum_{t=0}^{T}\lambda_t(s_0,a_0,\cdots,s_{t-1},a_{t-1},s_t),x_{T+1},s_{T+1})\right]
$$
$$
=\sup_{(a_t(s^t))\in\tilde{\mathcal{A}}^\infty(x_0)}\left[\mathbb{E}_{s_0}\sum_{t=0}^{T}\beta^t\left(r(x_t,a_t,s_t)\right.\right. \tag{A.16}
$$
$$
+\sum_{i=1}^{I}\left(\gamma^i+\sum_{n=0}^{t-1}\lambda_n^i(s_0,a_0,\cdots,s_{n-1},a_{n-1},s_n)\right)g^i(x_t,a_t(s^t),s_t)
$$
$$
\left.+\lambda_t^i(s_0,a_0,\cdots,s_{t-1},a_{t-1},s_t)\left(g^i(x_t,a_t,s_t)-\bar{g}^i\right)\right)
$$
$$
\left.+\mathbb{E}_{s_0}\beta^{T+1}F(\gamma+\sum_{t=0}^{T}\lambda_t(s_0,a_0,\cdots,s_{t-1},a_{t-1},s_t),x_{T+1},s_{T+1})\right]
$$

Since F is a fixed point of $\mathcal{B}$, we have

$$
\sup_{(a_t)\in\tilde{\mathcal{A}}^\infty(x_0)}\left[\mathbb{E}_{s_0}\beta^T\left(r(x_T,a_T,s_T)\right.\right.
$$
$$
+\sum_{i=1}^{I}\left(\gamma^i+\sum_{n=0}^{T-1}\lambda_n^i(s_0,a_0,\cdots,s_{n-1},a_{n-1},s_n)\right)g^i(x_T,a_T,s_T)
$$
$$
\left.+\lambda_T^i(s_0,a_0,\cdots,s_{T-1},a_{T-1},s_T)\left(g^i(x_T,a_T(s^T),s_T)-\bar{g}^i\right)\right)
$$
$$
\left.+\mathbb{E}_{s_0}\beta^{T+1}F(\gamma+\sum_{t=0}^{T}\lambda_t(s_0,a_0,\cdots,s_{t-1},a_{t-1},s_t),x_{T+1},s_{T+1})\right]
$$
$$
\geq\sup_{(a_t)\in\tilde{\mathcal{A}}^\infty(x_0)}\mathbb{E}_{s_0}\beta^T F(\gamma+\sum_{t=0}^{T-1}\lambda_t(s_0,a_0,\cdots,s_{t-1},a_{t-1},s_t),x_T,s_T).
$$

Hence by induction, the right hand side of (A.16) is bounded below by $F(\gamma,x_0,s_0)$, and (A.15) implies that $D(\gamma,x_0,s_0)+3\epsilon\geq F(\gamma,x_0,s_0)$. Since ϵ is arbitrarily chosen, we conclude that $D(\gamma,x_0,s_0)\geq F(\gamma,x_0,s_0)$. $\qquad\square$

Proof of Corollary 3.1

We define

$$
\tilde{\mathcal{B}}(F)=\sup_{a\in\tilde{\mathcal{A}}(x,s)}\inf_{\lambda\in\mathbb{R}_+^I}\left[\left(r(a,x,s)+\sum_{i=1}^{I}\gamma^i g^i(a,x,s)+\lambda^i(g^i(a,x,s)-\bar{g}^i)\right)+\beta\mathbb{E}_s F(\gamma+\lambda,x',s')\right].
$$

38

According to Lemma 3.1, we have

$$\tilde{\mathcal{B}}(F) = \sup_{a \in \tilde{\mathcal{A}}(x,s)} \inf_{\lambda \in \mathbb{R}_+^I} \left[\left(r(a,x,s) + \sum_{i=1}^I \gamma^i g^i(a,x,s) + \lambda^i(g^i(a,x,s) - \bar{g}^i) \right) + \beta \mathbb{E}_s F(\gamma + \lambda, x', s') \right]$$

$$= \inf_{(\lambda_a)_{a \in \mathcal{A}} \in \mathbb{R}_+^I} \sup_{a \in \tilde{\mathcal{A}}(x,s)} \left[\left(r(a,x,s) + \sum_{i=1}^I \gamma^i g^i(a,x,s) + \lambda_a^i(g^i(a,x,s) - \bar{g}^i) \right) + \beta \mathbb{E}_s F(\gamma + \lambda_a, x', s') \right].$$

To prove the corollary, it suffices to show that the operator $\tilde{\mathcal{B}}$ satisfies:

- $\tilde{\mathcal{B}}(N) \subset N$;
- $\tilde{\mathcal{B}}$ satisfies monotonicity,

for processing the same proof for Theorem 3.3 and Lemma 3.4. It is straightforward to verify the monotonicity. It suffices to check that $\tilde{\mathcal{B}}(N) \subset N$.

1. **Show that $\tilde{\mathcal{B}}$ preserves convexity.** We note that when F is convex in γ, then given $a \in \mathcal{A}$, $\left[\left(r(a,x,s) + \sum_{i=1}^I \gamma^i g^i(a,x,s) + \lambda_a^i(g^i(a,x,s) - \bar{g}^i) \right) + \beta \mathbb{E}_s F(\gamma + \lambda_a, x', s') \right]$ is convex in $(\gamma, (\lambda_a)_{a \in \mathcal{A}})$, then the same proof of Lemma 2 in Pavoni et al. (2018a) applies.

2. **Show that $\tilde{\mathcal{B}}$ preserves** (3.10), (3.12), **and the L-Lipschitz continuity.** The proof is analogous to that of Lemma 3.3, with λ replaced by λ_a.

3. **Show that $\tilde{\mathcal{B}}$ preserves the finiteness of the norm $\|\cdot\|_{\mathcal{M}}$.** This is a direct corollary of the fact that $\tilde{\mathcal{B}}(F)$ preserves the properties (3.10) and (3.12). $\qquad \square$

Proof of Theorem 4.1

- **Step 1.** For step 1, we aim to show that for any $x \in X$, we have

$$f^{**}(x) = \sup_{(x^*,\beta) \in X^* \times \mathbb{R}} \langle x^*, x \rangle - \beta,$$

$$\text{s.t.} \langle x^*, y \rangle - \beta \le f(y), \quad \forall y \in X. \tag{A.17}$$

First, by the definition of biconjugate, for any $x \in X$, we have

$$f^{**}(x) = \sup_{x^* \in X^*} \langle x^*, x \rangle - f^*(x^*)$$

$$= \sup_{(x^*,\beta) \in \text{epi}(f^*)} \langle x^*, x \rangle - \beta. \tag{A.18}$$

Second, by the definitions of epigraph and conjugate, we know that

$$(x^*, \beta) \in \text{epi}(f^*)$$

$$\Leftrightarrow f(x^*) \le \beta$$

$$\Leftrightarrow \sup_{y \in X} \langle x^*, y \rangle - f(y) \le \beta \tag{A.19}$$

$$\Leftrightarrow \langle x^*, y \rangle - f(y) \le \beta, \quad \forall y \in X$$

$$\Leftrightarrow \langle x^*, y \rangle - \beta \le f(y), \quad \forall y \in X.$$

39

Therefore (A.17) can be obtained by combining (A.18) and (A.19).

- **Step 2.** For step 2, we aim to show that

$$\text{epi}(f^{**}) \subseteq \text{cl co epi}(f). \tag{A.20}$$

For any $(x_0, \alpha_0) \notin \text{cl co epi}(f)$, according to Lemma B.1, there exists $x^* \in X$ and $\lambda \in \mathbb{R}$, such that

$$\langle x^*, x \rangle - \lambda \alpha < \gamma < \gamma_0 = \langle x^*, x_0 \rangle - \lambda \alpha_0, \quad \forall (x, \alpha) \in \text{cl co epi}(f). \tag{A.21}$$

Since f is proper, $\text{dom}(f) \neq \emptyset$, and hence there exists $x' \in \text{dom}(f)$. For any $\alpha' \geq f(x')$, $(x', \alpha') \in \text{epi}(f) \subseteq \text{cl co epi}(f)$. We take $x = x', \alpha = \alpha' \to +\infty$ on the left-hand-side in (A.21), and obtain $\lambda \geq 0$.

1. If $\lambda > 0$. According to (A.21), we know that

$$\langle \frac{x^*}{\lambda}, x \rangle - \frac{\gamma}{\lambda} < \alpha, \quad \forall (x, \alpha) \in \text{cl co epi}(f).$$

Therefore, $(\frac{x^*}{\lambda}, \frac{\gamma}{\lambda})$ satisfies the constraint in (A.17). Furthermore, from (A.21) we have

$$\langle \frac{x^*}{\lambda}, x_0 \rangle - \frac{\gamma}{\lambda} = \alpha_0 + \frac{\gamma_0 - \gamma}{\lambda} > \alpha_0,$$

implying that

$$f^{**}(x_0) > \alpha_0,$$

according to (A.17). Hence $(x_0, \alpha_0) \notin \text{epi}(f^{**})$.

2. If $\lambda = 0$, (A.21) implies that

$$\langle x^*, x \rangle < \gamma < \gamma_0 = \langle x^*, x_0 \rangle, \quad \forall x \in \text{epi}(f). \tag{A.22}$$

We take $\underline{x}^* \in X^*$, $\underline{\beta} \in \mathbb{R}$, such that (4.11) holds. Then for any $K > 0$, we have

$$\langle \underline{x}^* + Kx^*, x \rangle - (K\gamma + \beta)$$
$$= K(\langle x^*, x \rangle - \gamma) + \langle \underline{x}^*, x \rangle - \underline{\beta}$$
$$\leq \langle \underline{x}^*, x \rangle - \underline{\beta} \leq f(x), \quad \forall x \in X.$$

Therefore $(\underline{x}^* + Kx^*, K\gamma + \beta)$ satisfies the constraint in (A.17). Furthermore, from (A.22) we have

$$\langle \underline{x}^* + Kx^*, x_0 \rangle - (K\gamma + \beta)$$
$$= K(\langle x^*, x_0 \rangle - \gamma) + \langle \underline{x}^*, x_0 \rangle - \underline{\beta}$$
$$= K(\gamma_0 - \gamma) + \langle \underline{x}^*, x_0 \rangle - \beta,$$

implying that

$$f^{**}(x_0) \geq K(\gamma_0 - \gamma) + \langle \underline{x}^*, x_0 \rangle - \beta, \tag{A.23}$$

40

according to (A.17). We take

$$K > \frac{\alpha_0 + \beta - \langle \underline{x}^*, x_0 \rangle}{\gamma_0 - \gamma}$$

in (A.23) and obtain $f^{**}(x_0) > \alpha_0$, yielding that $(x_0, \alpha_0) \notin \mathrm{epi}(f^{**})$.

Therefore we have shown that, for any $(x_0, \alpha_0) \notin \mathrm{cl}\,\mathrm{co}\,\mathrm{epi}(f)$, we know that $(x_0, \alpha_0) \notin \mathrm{epi}(f^{**})$, which is equivalent to (A.20).

- **Step 3.** For step 3, we aim to show that

$$\mathrm{cl}\,\mathrm{co}\,\mathrm{epi}(f) \subseteq \mathrm{epi}(f^{**}). \tag{A.24}$$

From (A.17) we know $f^{**}(x) \leq f(x)$, $\forall x \in X$ from the constraint. Therefore, $\mathrm{epi}(f) \subseteq \mathrm{epi}(f^{**})$. It then suffices to show that $\mathrm{epi}(f^{**})$ is a closed convex set, or equivalently, to show that f^{**} is convex and lower semicontinuous, according to Proposition B.1 and Proposition B.2.

1. **Convexity.** Given $x \in X$. Assume that $x^* \in X^*$ satisfies $\langle x^*, y \rangle - \beta \leq f(y)$, $\forall y \in X$. According to (A.17), for any $x_1, x_2 \in X$, $0 \leq p_1, p_2 \leq 1$ such that $p_1 + p_2 = 1$ and $p_1 x_1 + p_2 x_2 = x$, we have

$$\begin{aligned}
&\langle x^*, x \rangle - \beta \\
=&\, p_1 \left(\langle x^*, y_1 \rangle - \beta \right) + p_2 \left(\langle x^*, y_2 \rangle - \beta \right) \\
\leq&\, p_1 f^{**}(y_1) + p_2 f^{**}(y_2).
\end{aligned} \tag{A.25}$$

Since (A.25) holds for any $x^* \in X^*$ satisfying $\langle x^*, y \rangle - \beta \leq f(y)$, $\forall y \in X$, we can deduce from (A.25) that

$$f^{**}(x) \leq p_1 f^{**}(x_1) + p_2 f^{**}(x_2).$$

Hence f^{**} is convex.

2. **Lower Semicontinuity.** To show that f^{**} is l.s.c., it suffices to show that for any $\epsilon > 0$ and any $\bar{x} \in X$, any sequence $\{x_n\}$ such that $x_n \to \bar{x}$ in X, we have

$$f^{**}(\bar{x}) - \epsilon \leq \liminf_{n \to \infty} f^{**}(x_n). \tag{A.26}$$

Indeed, according to (A.17), there exists $x^* \in X^*$ such that $\langle x^*, y \rangle - \beta \leq f(y)$, $\forall y \in X$, and

$$\langle x^*, \bar{x} \rangle - \beta \geq f^{**}(\bar{x}) - \frac{\epsilon}{2}. \tag{A.27}$$

Therefore, for any $x \in X$ such that

$$\|x - \bar{x}\|_X \leq \frac{\epsilon}{2 \max\{\|x^*\|_{X^*}, 1\}}, \tag{A.28}$$

we can deduce from (A.27) that

$$
\begin{aligned}
\langle x^*, x \rangle - \beta &= (\langle x^*, \bar{x} \rangle - \beta) + \langle x^*, x - \bar{x} \rangle \\
&\geq f^{**}(\bar{x}) - \frac{\epsilon}{2} + \|x^*\|_{X^*} \|x - \bar{x}\|_X \\
&\geq f^{**}(\bar{x}) - \frac{\epsilon}{2} - \frac{\epsilon}{2} \\
&= f^{**}(\bar{x}) - \epsilon.
\end{aligned}
\tag{A.29}
$$

Since (A.29) holds for any $x^* \in X^*$ satisfying $\langle x^*, y \rangle - \beta \leq f(y)$, $\forall y \in X$, we then know from (A.17) that

$$
f^{**}(x) \geq f^{**}(\bar{x}) - \epsilon,
$$

for all $x \in X$ satifying (A.28). And hence we obtain (A.26).

The proof is then finished by combining (A.20) in Step 2 and (A.24) in Step 3.

Proof of Theorem 4.2

- We first show that $p = v(\theta_Y)$.

  For any $x \in \Omega$ such that $g(x) \leq \theta_Y$, since $\langle y^*, g(x) \rangle \leq 0$, $\forall y^* \in Y_+^*$ and $\langle \theta_{Y^*}, g(x) \rangle = 0$, we have

$$
\sup_{y^* \in Y_+^*} L(x, y^*) = f(x).
\tag{A.30}
$$

For any $x \in \Omega$, such that $g(x) \leq \theta_Y$ does not hold, i.e. $g(x) \notin -P$, since $-P$ is a closed convex set, according to Lemma B.1, we know that there exists $y^* \in Y^*$ and $\gamma \in \mathbb{R}$, such that

$$
\langle y^*, y \rangle < \gamma < \langle y^*, g(x) \rangle, \quad \forall y \in -P.
\tag{A.31}
$$

Since $\theta_Y \in -P$ and $\langle y^*, \theta_Y \rangle = 0$, we know that $\gamma > 0$. On the other hand, if there exists $y \in -P$, such that $\langle y^*, y \rangle > 0$, then one can take $\alpha > 0$ sufficiently large, such that $\langle y^*, \alpha y \rangle = \alpha \langle y^*, y \rangle \geq \langle y^*, g(x) \rangle$, which contradicts (A.31). Therefore, $y^* \in Y_+^*$, and $\langle y^*, g(x) \rangle > \gamma > 0$. Hence for any $\alpha > 0$,

$$
\sup_{y^* \in Y_+^*} L(x, y^*) \geq f(x) + \alpha \langle y^*, g(x) \rangle > f(x) + \alpha \gamma,
$$

implying that

$$
\sup_{y^* \in Y_+^*} L(x, y^*) = +\infty.
\tag{A.32}
$$

We combine (A.30) and (A.32) to obtian

$$
\sup_{y^* \in Y_+^*} L(x, y^*) = \begin{cases} f(x), & g(x) \leq \theta_Y; \\ +\infty, & g(x) \not\leq \theta_Y, \end{cases}
$$

yielding that $p = \inf_{x \in \Omega} \sup_{y^* \in Y_+^*} L(x, y^*) = v(\theta_Y)$.

- For step 2, we show that $d = v^{**}(\theta_Y)$. By direct compuataion, we have

$$
\begin{aligned}
v^{**}(\theta_Y) &= \sup_{y^* \in Y^*} -v^*(y^*) \\
&= \sup_{y^* \in Y^*} - \left(\sup_{y \in Y} (\langle y^*, y \rangle - v(y)) \right) \\
&= \sup_{y^* \in Y^*} \inf_{y \in Y} (v(y) - \langle y^*, y \rangle) \\
&= \sup_{y^* \in Y^*} \inf_{y \in Y} (\inf_{g(x) \le y} f(x) - \langle y^*, y \rangle) \\
&= \sup_{y^* \in Y^*} \inf_{x \in \Omega} \inf_{y \ge g(x)} (f(x) - \langle y^*, y \rangle) \\
&= \sup_{y^* \in Y^*} \inf_{x \in \Omega} \left[(f(x) - \langle y^*, g(x) \rangle) + \inf_{y \ge g(x)} \langle y^*, g(x) - y \rangle \right] \\
&= \sup_{y^* \in Y^*} \inf_{x \in \Omega} \left[(f(x) - \langle y^*, g(x) \rangle) + \inf_{y \ge g(x)} \langle y^*, y - g(x) \rangle \right].
\end{aligned}
\tag{A.33}
$$

If $y^* \notin Y_+^*$, then it is straightforward to verify that for any $x \in \Omega$, there exists $y \ge g(x)$, such that

$$
\inf_{y \ge g(x)} \langle y^*, y - g(x) \rangle = -\infty,
$$

and hence

$$
\inf_{x \in \Omega} \left[(f(x) + \langle y^*, g(x) \rangle) + \inf_{y \ge g(x)} \langle y^*, y - g(x) \rangle \right] = -\infty.
$$

If $y^* \in Y_+^*$, then for any $x \in \Omega$, $y \ge g(x)$, we have

$$
\langle y^*, y - g(x) \rangle \ge 0,
$$

and the equality is achieved when $y = g(x)$. Therefore, we conclude from (A.33) that

$$
v^{**}(\theta_Y) = \sup_{y^* \in Y_+^*} \inf_{x \in \Omega} (f(x) + \langle y^*, g(x) \rangle) = \sup_{y^* \in Y_+^*} \inf_{x \in \Omega} L(x, y^*) = d.
$$

We combine the two steps to finish the proof of the first statement. The fact that $\mathrm{epi}(v^{**}) = \mathrm{cl} \, \mathrm{co} \, \mathrm{epi}(v)$ follows directly from Theorem 4.1. $\qquad \square$

Proof of Theorem 4.3

We consider a sequence of $P_n \in \mathcal{P}(\tilde{\mathcal{A}}^\infty(x_0))$ satisfies all constraints in (4.12), and

$$
\mathbb{E}_{s_0}^{(a(s^t)) \sim P_n} \sum_{t=0}^\infty \beta^t \left(r(x(s^t, a^{t-1}), a(s^t), s_t) + \sum_{i=1}^I \gamma^i g^i(x(s^t, a^{t-1}), a(s^t), s_t) \right)
$$

tends to the supremum of problem (4.12). According to the *-weak compactness of $\mathcal{P}(\tilde{\mathcal{A}}^\infty(x_0))$, P_n *-weak converges to some $P^* \in \mathcal{P}(\tilde{\mathcal{A}}^\infty(x_0))$ up to a subsequence.

For any $a \in \tilde{A}^\infty(x_0)$, we define

$$f(a) = \mathbb{E}_{s_0} \sum_{t=0}^\infty \beta^t \left(r(x(s^t, a^{t-1}), a(s^t), s_t) + \sum_{i=1}^I \gamma^i g^i(x(s^t, a^{t-1}), a(s^t), s_t) \right),$$

and

$$g_{t,i,h^t}(a) = \mathbb{E}_{s^t} 1_{\{\tilde{a}^{t-1} = a^{t-1}\}} \left(\sum_{n=0}^\infty \beta^n g^i(x(s^{t+n}, (\tilde{a}^{t-1}, a_t^{t+n-1})), a(s^{t+n}), s_{t+n}) - \bar{g}^i \right).$$

It is straightforward to see that f and g_{t,i,h^t} are bounded, and continuous in a under the product topology. Hence $\mathbb{E}^{a \sim P} f(a) \to \mathbb{E}^{a \sim P^*} f(a)$, and $\mathbb{E}^{a \sim P} g_{t,i,h^t}(a) \to \mathbb{E}^{a \sim P^*} g_{t,i,h^t}(a)$, implies that P^* is a feasible point to the problem (4.12) and reaches the supremum.

Proof of Theorem 4.4

For any $P \in \mathcal{P}(\tilde{\mathcal{A}}^\infty(x_0))$ which is a feasible point to problem (4.12), it is straightforward to compute $\psi(h^t)(a)$ as[13]

$$\psi(h^t)(a) = \frac{P(a(s^t) = a, a(s^{t-1}) = \tilde{a}_{t-1}, \cdots, a(s_0) = \tilde{a}_0)}{P(a(s^{t-1}) = \tilde{a}_{t-1}, \cdots, a(s_0) = \tilde{a}_0)}.$$

It is straightforward to show that $(\psi(h^t))$ is a feasible point to (4.13), and reaches the same objective function value as P in problem (4.12).

For $(\psi(h^t)) \in \Pi_{h^t \in \mathcal{H}^t} \mathcal{P}(\tilde{A}(h^t))$, we can construct the marginal distributions as

$$\psi(a(s^t) = a, a(s^{t-1}) = \tilde{a}_{t-1}, \cdots, a(s_0) = \tilde{a}_0) = \psi(h_0)(\tilde{a}_0) \cdot \psi(h^1)(\tilde{a}_1) \cdots \psi(h^t)(a).$$

According to Kolmogorov extension theorem, there exists $P \in \mathcal{P}(\tilde{\mathcal{A}}^\infty(x_0))$ has marginal distributions ψ, and hence P reaches the same objective function value as $(\psi(h^t))$ in problem (4.13). □

Proof of Theorem 4.5

According to Theorem C.3, $D(\gamma, x_0, s_0)$ equals to the dual of the deterministic problem (4.6). According to Corollary B.2, for any $\epsilon > 0$, there exists $N \in \mathbb{N}_+$, $(a_t^k) \in \tilde{\mathcal{A}}^\infty(x_0)$ $(k = 1, \cdots, N)$, $p_k \geq 0$, $\sum_{k=1}^N p_k = 1$, and $P_\epsilon \in \mathcal{P}(\tilde{\mathcal{A}}^\infty(x_0))$ defined by

$$P_\epsilon(\{(a_t^k)\}) = p_k,$$

such that

$$\mathbb{E}_{s_0}^{(a_t(s^t)) \sim P_\epsilon} \sum_{t=0}^\infty \beta^t \left(r(x_t, a_t, s_t) + \sum_{i=1}^I \gamma^i g^i(x_t, a_t, s_t) \right) \geq D(\gamma, x_0, s_0) - \epsilon,$$

─────────────

[13]Note that when $P(a(s^{t-1}) = \tilde{a}_{t-1}, \cdots, a(s_0) = \tilde{a}_0) = 0$ we can arbitrarily define $\psi(h^t)$.

and

$$\mathbb{E}_{s^t}^{(a(s^t))\sim P_\epsilon} 1_{\{(\tilde{a}_0,\cdots,\tilde{a}_{t-1})=(a(s_0),\cdots,a(s^{t-1}))\}} \left(\sum_{n=0}^{\infty} \beta^n g^i(x_{t+n}, a(s^{t+n}), s_{t+n}) - \bar{g}^i \right) \geq -\epsilon,$$

$$\forall t \in \mathbb{N},\ \forall h^t = (s_0, \tilde{a}_0, \cdots, s_{t-1}, \tilde{a}_{t-1}, s_t) \in \mathcal{H}^t,\ \forall i \in \{1, \cdots, I\}.$$

Since $\mathcal{P}(\tilde{\mathcal{A}}^\infty(x_0))$ is *-weak compact, P_ϵ *-weak converges to some $P^* \in \mathcal{P}(\tilde{\mathcal{A}}^\infty(x_0))$ as $\epsilon \to 0$ up to a subsequence. It is straightforward to verify that P^* is a feasible probability measure to problem (4.12) and

$$\mathbb{E}_{s_0}^{(a(s^t))\sim P^*} \sum_{t=0}^{\infty} \beta^t \left(r(x_t, a(s^t), s_t) + \sum_{i=1}^{I} \gamma^i g^i(x_t, a(s^t), s_t) \right) \geq D(\gamma, x_0, s_0).$$

Therefore, $V(\gamma, x_0, s_0) \geq D(\gamma, x_0, s_0)$.

Then it suffices to prove that $V(\gamma, x_0, s_0) \leq D(\gamma, x_0, s_0)$. We consider the set of finite support probability measures,

$$\mathcal{D} = \{\sum_{i=1}^{n} \lambda_i \delta_{x_i} | n \in \mathbb{N},\ \lambda_i \geq 0,\ \sum_{i=1}^{n} \lambda_i = 1,\ x_i \in \tilde{\mathcal{A}}^\infty(x_0)\}.$$

$\mathcal{D}$ is *-weakly dense in $\mathcal{P}(\tilde{\mathcal{A}}^\infty(x_0))$(see Theorem 15.10 in Aliprantis and Border (2006)), i.e. for any $P^* \in \mathcal{P}(\tilde{\mathcal{A}}^\infty(x_0))$, there exists $\{P_D^k\}_{k=1}^{\infty} \subset \mathcal{D}$, s.t.

$$\int_{a \in \tilde{\mathcal{A}}^\infty(x_0)} f(a) dP_D^k \to \int_{a \in \tilde{\mathcal{A}}^\infty(x_0)} f(a) dP^*,$$

for any $f \in C_b(\tilde{\mathcal{A}}^\infty(x_0))$. We take $P^* \in \mathcal{P}(\tilde{\mathcal{A}}^\infty(x_0))$ as the solution to (4.12), and denote

$$P_D^k = \sum_{i=1}^{n_k} \lambda_{k,i} \delta_{a_{k,i}}.$$

According to the fifth assumption in Assumption 3.1, for any $a \in \tilde{\mathcal{A}}^\infty(x_0)$, there exists $(\underline{a}_t(s^t)) \in \mathcal{A}^\infty$, such that when defining $a^T \in \mathcal{A}^\infty$ as

$$a^T = \begin{cases} a_t(s^t), & t \leq T; \\ \underline{a}_t(s^t), & t > T, \end{cases}$$

$a^T \in \tilde{\mathcal{A}}^\infty(x_0)$ and satisfies the constraints in (3.1) for any $t > T$. Furthermore, we can define

$$\underline{P}_D^k = \sum_{i=1}^{n_k} \lambda_{k,i} \delta_{a_{k,i}^k}.$$

To finish the proof two lemmas are needed.

Lemma A.1 *For any $P \in \mathcal{P}(\tilde{\mathcal{A}}^\infty(x_0))$, we define*

$$G(P) =$$

$$(\min\{\mathbb{E}_{s^t}^{(a(s^t)) \sim P} 1_{\{\tilde{a}^{t-1} = a^{t-1}\}} \left(\sum_{n=0}^{\infty} \beta^n g^i(x_{t+n}, a(s^{t+n}), s_{t+n}) - \bar{g}^i \right), 0\})_{t \in \mathbb{N}, \, h^t \in \mathcal{H}^t, \, i \in \{1, \cdots, I\}} \in \ell^\infty.$$

Then $G(\underline{P}_D^k)$ weak converges to θ in ℓ^∞.

Proof By definition, it suffices to show that for any $\lambda^1 \in \ell^1$, $\lambda^s \in \ell^s$, we have

$$\langle \lambda^1 + \lambda^s, G(\underline{P}_D^k) \rangle \to 0, \quad \text{as } k \to \infty. \tag{A.34}$$

By the definition of $\underline{P}_D^k$, it is easy to check that

$$G(\underline{P}_D^k)_{t > n_k, \, h^t \in \mathcal{H}^t, \, i \in \{1, \cdots, I\}} = 0,$$

yielding that

$$\langle \lambda^s, G(\underline{P}_D^k) \rangle = 0, \quad \forall k \in \mathbb{N}_+. \tag{A.35}$$

For any t, h^t, i, for $a \in \tilde{\mathcal{A}}^\infty(x_0)$, we define

$$g_{t,h^t,i}(a) := \mathbb{E}_{s_t} 1_{\{(\tilde{a}^{t-1} = a^{t-1})\}} \left(\sum_{n=0}^{\infty} \beta^n g^i(x_{t+n}, a(s^{t+n}), s_{t+n}) - \bar{g}^i \right).$$

Therefore,

$$\int_{a \in \tilde{\mathcal{A}}^\infty(x_0)} g_{t,h^i,i}(a) dP_D^k \to \int_{a \in \tilde{\mathcal{A}}^\infty(x_0)} g_{t,h^t,i}(a) dP^* \geq 0, \quad \text{as } k \to \infty, \quad \forall t, h^t, i,$$

due to the fact that P_D^k weak converges to P^*. Hence

$$G(P_D^k)_{t,h^t,i} = \min\{ \int_{a \in \tilde{\mathcal{A}}^\infty(x_0)} g_{t,h^t,i}(a) dP_D^k, 0\} \to 0, \quad \text{as } k \to \infty, \quad \forall t, h^t, i. \tag{A.36}$$

It is straightforward to verify that[14]

$$G(\underline{P}_D^k)_{t,h^t,i} - G(P_D^k)_{t,h^t,i} \to 0, \quad \text{as } k \to \infty, \quad \forall t, h^t, i. \tag{A.37}$$

We combine (A.36) and (A.37) to obtain

$$G(\underline{P}_D^k)_{t,h^t,i} \to 0.$$

This pointwise convergence property, together with the fact that $G(\underline{P}_D^k)$ is uniformly bounded

---

[14]This is the UANA property.

in ℓ^∞, yields that $G(\underline{P}_D^k)$ *-weak converges to θ in ℓ^∞, implying that

$$\langle \lambda^1, G(\underline{P}_D^k) \rangle \to 0, \text{ as } k \to \infty. \tag{A.38}$$

We then conclude (A.34) by combining (A.35) and (A.38). $\square$ $\square$

Lemma A.2 *For any $P \in \mathcal{P}(\tilde{\mathcal{A}}^\infty(x_0))$, we define*

$$F(P) = \mathbb{E}_{s_0}^{(a(s^t)) \sim P} \sum_{t=0}^\infty \beta^t \left(r(x_t, a(s^t), s_t) + \sum_{i=1}^I \gamma^i g^i(x_t, a(s^t), s_t) \right).$$

Then $F(\underline{P}_D^k)$ converges to $F(P^*) = V(\gamma, x_0, s_0)$.

Proof For $a \in \tilde{\mathcal{A}}^\infty(x_0)$, we define

$$f(a) := \mathbb{E}_{s_0} \sum_{t=0}^\infty \beta^t \left(r(x_t, a(s^t), s_t) + \sum_{i=1}^I \gamma^i g^i(x_t, a(s^t), s_t) \right).$$

Therefore,

$$F(P_D^k) = \int_{a \in \tilde{\mathcal{A}}^\infty(x_0)} f(a) dP_D^k \to \int_{a \in \tilde{\mathcal{A}}^\infty(x_0)} f(a) dP^* = F(P^*), \quad \text{as } k \to \infty.$$

Similarly as in the proof for Lemma A.1, it is straightforward to verify that

$$F(\underline{P}_D^k) - F(P_D^k) \to 0, \quad \text{as } k \to \infty,$$

and we omit details here. $\square$

Let v be the perturbation functional of the deterministic problem (3.1) [15]. It can be verified by definition that

$$(G(\underline{P}_D^k), F(\underline{P}_D^k)) \in \text{co hypo } (v).$$

According to Mazur's Theorem(see Corollary 3.8 in Brezis and Brézis (2011)), the weak limit of $(G(\underline{P}_D^k), F(\underline{P}_D^k))$ belongs to cl co hypo (v), i.e.

$$(\theta, V(\gamma, x_0, s_0)) \in \text{cl co hypo } (v) = \text{hypo } (v^{**}),$$

where the last equality holds according to Theorem 4.1. According to Theorem 4.2,

$$D(\gamma, x_0, s_0) = v^{**}(\theta).$$

Therefore $V(\gamma, x_0, s_0) \leq D(\gamma, x_0, s_0)$ by definition of hypograph.

---

[15]We refer to Definition B.3 for the defintion of a perturbation functional for a minimization problem. Note that this problem is indeed a maximization problem, which is different from the minimization problem considered in standard optimization literatures. Hence the concept epigraph is replaced by hypograph, which can be similarly defined and we omit details here.

Proof of Corollary 4.1

It suffices to show that $V \in \mathcal{N}$. We denote $P^*(\gamma, x_0, s_0)$ to be a maximizer for any $\gamma \in \mathbb{R}_+^I$, $x_0 \in \mathcal{X}$, $s_0 \in \mathcal{S}$.

- **Convexity**. For any γ_0, $\gamma_1 \in \mathbb{R}_+^I$, and $\lambda \in [0,1]$, we denote $\gamma_\lambda = \lambda\gamma_1 + (1-\lambda)\gamma_0$. Then $P^*(\gamma_\lambda, x_0, s_0)$ is a feasible point for the problem (4.12) with x_0, s_0. Therefore,

$$
\lambda V(\gamma_1, x_0, s_0) + (1-\lambda)V(\gamma_0, x_0, s_0)
$$
$$
\geq \lambda \mathbb{E}_{s_0}^{(a(s^t)) \sim P^*(\gamma_\lambda, x_0, s_0)} \sum_{t=0}^{\infty} \beta^t \left(r(x_t, a(s^t), s_t) + \sum_{i=1}^{I} \gamma^i g^i(x_t, a(s^t), s_t) \right)
$$
$$
+ (1-\lambda) \mathbb{E}_{s_0}^{(a(s^t)) \sim P^*(\gamma_\lambda, x_0, s_0)} \sum_{t=0}^{\infty} \beta^t \left(r(x_t, a(s^t), s_t) + \sum_{i=1}^{I} \gamma^i g^i(x_t, a(s^t), s_t) \right)
$$
$$
= V(\gamma_\lambda, x_0, s_0).
$$

- *L*-**Lipschitz** For any γ_0, $\gamma_1 \in \mathbb{R}_+^I$, we have

$$
V(\gamma_1, x_0, s_0) \geq V(\gamma_0, x_0, s_0) + \mathbb{E}_{s_0}^{(a(s^t)) \sim P^*(\gamma_0, x_0, s_0)} \sum_{t=0}^{\infty} \beta^t \sum_{i=1}^{I} (\gamma_1^i - \gamma_0^i) g^i(x_t, a(s^t), s_t)
$$
$$
\geq V(\gamma_0, x_0, s_0) - \frac{\|\gamma_1 - \gamma_0\|_1 \max_i \|g^i\|_\infty}{1 - \beta}
$$
$$
\geq V(\gamma_0, x_0, s_0) - L\|\gamma_1 - \gamma_0\|_1
$$

By symmetric, we have

$$
V(\gamma_0, x_0, s_0) \geq V(\gamma_1, x_0, s_0) - L\|\gamma_1 - \gamma_0\|_1.
$$

- **The bound** (3.10) **and** (3.12). (3.10) is obtianed by definition of V. (3.12) is obtained by the upper bounded of r and g^i. $\qquad\square$

Proof of Theorem 4.6

We consider the problem

$$
\sup_{(a(s^t)) \in \tilde{\mathcal{A}}^\infty(x_0) \subset \ell^\infty} \mathbb{E}_{s_0} \sum_{t=0}^{\infty} \beta^t \left(r(x(s^t), a(s^t), s_t) + \sum_{i=1}^{I} \gamma^i g^i(x(s^t), a(s^t), s_t) \right),
$$
$$
\text{s.t. } 1_{\{(\tilde{a}_0, \cdots, \tilde{a}_t) = (a(s_0), \cdots, a(s^t))\}} \left(\mathbb{E}_{s_t} \sum_{n=0}^{\infty} \beta^n g^i(x(s^{t+n}), a(s^{t+n}), s_{t+n}) - \bar{g}^i \right) \geq 0,
$$
$$
\forall t \in \mathbb{N}, \forall \tilde{h}^t = (s_0, \tilde{a}_0, \cdots, s_t, \tilde{a}_t) \in \tilde{\mathcal{H}}^t, \forall i \in \{1, \cdots, I\}.
$$
(A.39)

Following reasoning analogous to that in Theorem C.3, we find that $\tilde{D}(\gamma, x_0, s_0)$ equals the Lagrangian dual value of problem (A.39). Then, employing a similar line of argument as in Theorem 4.5 and Corollary 4.1, we arrive at the conclusion of this theorem.

48

Proof of Theorem 5.1

Since (3.5) is convex w.r.t. λ, to prove the existence of a solution λ^*, it suffices to show that, there exists $C > 0$, s.t. when $\lambda_i > C$ for some i,

$$\sup_{a \in \mathcal{A}} \left[r(x, a, s) + \sum_{i=1}^{I} \gamma^i g^i(x, a, s) + \lambda^i(g^i(x, a, s) - \bar{g}^i) + \beta \mathbb{E}_s D(\gamma + \lambda, x', s') \right] > D(\gamma, x, s). \tag{A.40}$$

According to the property (3.10), we know that the LHS of (5.1) is not less than

$$\mathbb{E}_{s_0}^{a(h^t) \sim \psi(h^t)} \sum_{t=0}^{\infty} \beta^t r(x_t, a_t, s_t) + \sum_{i=1}^{I} \gamma^i \mathbb{E}_{s_0}^{a(h^t) \sim \psi(h^t)} \sum_{t=0}^{\infty} \beta^t g^i(x_t, a_t, s_t) + \sum_{i=1}^{I} \lambda^i \mathbb{E}_{s_0}^{a(h^t) \sim \psi(h^t)} \left(\sum_{t=0}^{\infty} \beta^t g^i(x_t, a_t, s_t) - \bar{g}^i \right)$$

$$\geq - \left(1 + \sum_{i=1}^{I} \gamma^i \right) L + \sum_{i=1}^{I} \lambda^i \epsilon$$

It is then straightforward to verify that it follows from Assumption 5.1 that, taking

$$C = 2 \frac{D(\gamma, x, s) + \left(1 + \sum_{i=1}^{I} \gamma^i \right) L}{\epsilon},$$

when $\lambda_i > C$, (A.40) is satisfied.

Proof of Lemma 5.1

We define the set

$$\mathcal{U} = \{(u, v) \in \mathbb{R}^{I+1} | \exists P \in \mathcal{P}(\tilde{\mathcal{A}}^{\infty}(x_0)), P \text{ satisfies all constraints in (4.12) for } x_0 = x, s_0 = s;$$

$$u = \mathbb{E}_{s_0}^{(a_t \sim P)} \sum_{t=0}^{\infty} \beta^t r(x_t, a_t, s_t);$$

$$v = \mathbb{E}_{s_0}^{(a_t \sim P)} \sum_{t=0}^{\infty} \beta^t g(x_t, a_t, s_t) \}.$$

It is straightforward to see that $\mathcal{U}$ is a non-empty convex set. Furthermore, based on the compactness of $\mathcal{P}(\tilde{\mathcal{A}}^{\infty}(x_0))$ in the *-weak topology, we know that $\mathcal{U}$ is closed. We define

$$F(\gamma_1, \gamma) = \sup_{(u,v) \in \mathcal{U}} \gamma_1 u + \gamma \cdot v.$$

According to Corollary 23.5.3 in Rockafellar (2015), we know that $(u^*, v^*) \in \arg\sup_{(u,v) \in \mathcal{U}} \gamma_1 u + \gamma \cdot v \Leftrightarrow (u^*, v^*) \in \partial F(\gamma_1, \gamma)$. The dual function equals

$$D(\gamma, x, s) = \sup_{(u,v) \in \mathcal{U}} u + \gamma \cdot v.$$

We take $A = [0, I_I]$, $b = [1, 0_{I \times 1}]$, and have

$$D(\gamma, x, s) = F(1, \gamma) = F(A\gamma + b).$$

According to Theorem 23.9 in Rockafellar (2015), we know that

$$\partial D(\gamma, x, s) = A^T \partial F(A\gamma + b).$$

Therefore,

$$
\begin{aligned}
&\phi \in \partial D(\gamma, x, s) \\
\Leftrightarrow &\phi = A^T \phi_F, \text{ for some } \phi_F \in \partial F(1, \gamma) \\
\Leftrightarrow &\phi = (\phi_F(2), \phi_F(3), \cdots, \phi_F(I+1)), \text{ for some } (\phi_F(1), \phi_F(2:I+1)) \in \arg \sup_{(u,v) \in \mathcal{U}} u + \gamma v \\
\Leftrightarrow &\exists P^* \in \mathcal{P}(\tilde{\mathcal{A}}^\infty(x_0)), P^* \text{ maximizes problem (4.12), and } \mathbb{E}_{s_0}^{(a \sim P)} \sum_{t=0}^\infty \beta^t g(x_t, a_t, s_t) = \phi.
\end{aligned}
$$

Hence we finish the proof.

Proof of Theorem 5.3

Since $\phi \in \partial D(\gamma, x, s)$, we know that there exists $(\psi(h^t))$, s.t. $(\psi(h^t))$ is a solution to (5.1) with

$$
\begin{aligned}
\phi =& \mathbb{E}_{s_0}^{a(h^t) \sim \psi(h^t)} \sum_{t=0}^\infty \beta^t g(x_t, a_t, s_t) \\
=& \sum_{a \in \mathcal{A}} \psi_0(a) \left[g(x_0, a, s_0) + \beta \mathbb{E}_{s_0} \mathbb{E}_{h^1 = (s_0, a, s_1)}^{a(h^t) \sim \psi(h^t)} \sum_{t=1}^\infty \beta^{t-1} g(x_t, a_t, s_t) \right].
\end{aligned}
$$

We take $\psi^* = \psi_0$ and $\phi'^*(a, s_1) = \mathbb{E}_{h^1 = (s_0, a, s_1)}^{a(h^t) \sim \psi(h^t)} \sum_{t=1}^\infty \beta^{t-1} g(x_t, a_t, s_t)$. Hence (5.5) is satisfied.

It is straightforward to verify that, for any λ^* be a solution to (3.5), and $\mu^* = 0$, we have

$$
\begin{aligned}
& \sum_{a \in \mathcal{A}} \psi^*(a) \\
& \left[\left(r(x, a, s) + \sum_{i=1}^I \gamma^i g^i(x, a, s) + \lambda^{*,i}(g^i(x, a, s) - \bar{g}^i) + \mu^{*,i}(g^i(x, a, s) - \phi^i) \right) + \beta \mathbb{E}_s W(\gamma + \lambda^* + \mu^*, x', s', \phi'^*(a, s')) \right] \\
\geq & \mathbb{E}_{s_0}^{(a(h^t) \sim \psi(h^t))} \left[\sum_{t=0}^\infty \beta^t r(x_t, a_t, s_t) + \sum_{t=0}^\infty \beta^t \gamma g(x_t, a_t, s_t) + \lambda^*(-\bar{g} + \sum_{t=0}^\infty \beta^t g(x_t, a_t, s_t)) \right] \\
= & \mathbb{E}_{s_0}^{(a(h^t) \sim \psi(h^t))} \left[\sum_{t=0}^\infty \beta^t r(x_t, a_t, s_t) + \sum_{t=0}^\infty \beta^t \gamma g(x_t, a_t, s_t) \right] + \lambda^* \left(\sum_{a \in \mathcal{A}} \psi(a) \left(g(x, a, s) + \beta \mathbb{E}_s \phi'(a, s') - \bar{g} \right) \right) \\
\geq & \mathbb{E}_{s_0}^{(a(h^t) \sim \psi(h^t))} \left[\sum_{t=0}^\infty \beta^t r(x_t, a_t, s_t) + \sum_{t=0}^\infty \beta^t \gamma g(x_t, a_t, s_t) \right] = W(\gamma, x, s, \phi),
\end{aligned}
$$

with equality only if (5.6) is satisfied. Furthermore, ψ^*, ϕ'^* should be a solution to the sup problem given $\lambda^*, \mu^* = 0$, and $(\psi(h^t|s_0, a, s')(t \geq 1))$ is a solution to $W(\gamma + \lambda^*, \mu^*, x', s', \phi'^*(a, s'))$, hence (5.4) shoule be satisfied according to lemma 5.1. Therefore $(\lambda^*, 0, \psi^*, \phi'^*)$ is a solution to (5.3) satisfying (5.4), (5.5) and (5.6) □

Proof of Theorem 5.4

From the ≥ 0 part of (5.5) and (5.6), and the fact that $\phi^*(h^t)$ are bounded, one can verify directly that $(\psi^*(h^t))$ satisfies all the constraints. Furthermore, we know that

$$W(\gamma_0, x_0, s_0, \phi_0)$$

$$= \sum_{a_0 \in \mathcal{A}} \psi_0^*(a_0) \Big[r(x_0, a_0, s_0) + \gamma_0 g(x_0, a_0, s_0) + \lambda_0^*(g(x_0, a_0, s_0) - \bar{g}) + \mu_0^*(g(x_0, a_0, s_0) - \phi_0)$$

$$+ \beta \mathbb{E}_{s_0} W(\gamma_0 + \lambda_0^* + \mu_0^*, x_1, s_1, \phi_1^*) \Big]$$

$$= \sum_{a_0 \in \mathcal{A}} \psi_0^*(a_0) \Big\{ r(x_0, a_0, s_0) + \gamma_0 g(x_0, a_0, s_0) + \lambda_0^*(g(x_0, a_0, s_0) - \bar{g}) + \mu_0^*(g(x_0, a_0, s_0) - \phi_0)$$

$$+ \beta \mathbb{E}_{s_0} \Big[\sum_{a_1 \in \mathcal{A}} \psi_1^*(a_1) \Big(r(x_1, a_1, s_1) + (\gamma_0 + \lambda_0^* + \mu_0^*)g(x_1, a_1, s_1) + \lambda_1^*(g(x_1, a_1, s_1) - \bar{g}) + \mu_1^*(g(x_1, a_1, s_1) - \phi_1)$$

$$\beta^2 \mathbb{E}_{s_1} W(\gamma_0 + \lambda_0^* + \lambda_1^* + \mu_0^* + \mu_1^*, x_2, s_2, \phi_2^*) \Big) \Big] \Big\}$$

$$= \cdots$$

$$= \mathbb{E}_{s_0}^{(a(h^t)) \sim (\psi(h^t))} \Big[\sum_{t=0}^{T} \beta^t r(x_t, a_t, s_t) + \gamma_0 \sum_{t=0}^{T} \beta^t g(x_t, a_t, s_t) + \sum_{t=0}^{T} \sum_{h^t \in \mathcal{H}^t} \lambda^*(h^t) \Big(-\beta^t \bar{g} + \sum_{r=t}^{T} \beta^r g(x_r, a_r, s_r) \Big)$$

$$+ \sum_{t=0}^{T} \sum_{h^t \in \mathcal{H}^t} \mu^*(h^t) \Big(-\beta^t \phi_t^* + \sum_{r=t}^{T} \beta^r g(x_r, a_r, s_r) \Big) + \beta^{T+1} W(\gamma_{T+1}^*, x_{T+1}, s_{T+1}, \phi_{T+1}^*) \Big]$$

Note that $\phi_{T+1}^* \in \partial D(\gamma_{T+1}^*, x_{T+1}, s_{T+1})$, according to Lemma 5.1,

$$W(\gamma_{T+1}^*, x_{T+1}, s_{T+1}, \phi_{T+1}^*) = u_{T+1}^* + \gamma_{T+1}^* \phi_{T+1}^*,$$

for some $u_{T+1}^* \leq L$.

Hence

$$W(\gamma_0, x_0, s_0, \phi_0)$$

$$=\mathbb{E}_{s_0}^{(a(h^t))\sim(\psi(h^t))}\left[\sum_{t=0}^{T}\beta^t r(x_t, a_t, s_t) + \gamma_0\sum_{t=0}^{T}\beta^t g(x_t, a_t, s_t) + \sum_{t=0}^{T}\sum_{h^t\in\mathcal{H}^t}\lambda^*(h^t)\left(-\beta^t\bar{g} + \sum_{r=t}^{T}\beta^r g(x_r, a_r, s_r)\right)\right.$$

$$+\sum_{t=0}^{T}\sum_{h^t\in\mathcal{H}^t}\mu^*(h^t)\left(-\beta^t\phi^* + \sum_{r=t}^{T}\beta^r g(x_r, a_r, s_r)\right) + \left.\beta^{T+1}W(\gamma_{T+1}^*, x_{T+1}, s_{T+1}, \phi_{T+1}^*)\right]$$

$$=\mathbb{E}_{s_0}^{(a(h^t))\sim(\psi(h^t))}\left[\sum_{t=0}^{T}\beta^t r(x_t, a_t, s_t) + \beta^{T+1}u_{T+1}^* + \gamma_0\left(\sum_{t=0}^{T}\beta^t g(x_t, a_t, s_t) + \beta^{T+1}\phi_{T+1}'^*\right)\right.$$

$$+\sum_{r=0}^{T}\sum_{h^t\in\mathcal{H}^t}\lambda^*(h^t)\left(-\beta^t\bar{g} + \sum_{r=t}^{T}\beta^r g(x_r, a_r, s_r) + \beta^{T+1}\phi_{T+1}'^*\right)$$

$$+\sum_{t=0}^{T}\sum_{h^t\in\mathcal{H}^t}\mu^*(h^t)\left(-\beta^t\phi_t^* + \sum_{r=t}^{T}\beta^r g(x_r, a_r, s_r) + \left.\beta^{T+1}\phi_{T+1}'^*\right)\right]$$

$$=\mathbb{E}_{s_0}^{(a(h^t))\sim(\psi(h^t))}\left[\sum_{t=0}^{T}\beta^t r(x_t, a_t, s_t) + \beta^{T+1}u_{T+1}^* + \gamma_0\sum_{t=0}^{T}g(x_t, a_t, s_t) + \beta^{T+1}\phi_{T+1}'^*\right],$$

according to the complementary conditions (5.5) and (5.6). Since

$$\beta^{T+1}u_{T+1}^* + \gamma_0\beta^{T+1}\phi_{T+1}^*$$

uniformly converges to 0, we konw that

$$W(\gamma_0, x_0, s_0, \phi_0) = \mathbb{E}_{s_0}^{(a_t(h^t))\sim(\psi_t(h^t))}\left(\sum_{t=0}^{\infty}\beta^t r(x_t, a_t, s_t) + \gamma_0\sum_{t=0}^{\infty}\beta^t g(x_t, a_t, s_t)\right)$$

and hence finish the proof. □

Proof of Theorem 5.5

Step 1. Existence. Since $\phi \in \partial D(\gamma, x, s)$, we know that there exists $(\psi(h^t))$, s.t. $(\psi(h^t))$ is a solution to (5.1) with

$$\phi =\mathbb{E}_{s_0}^{a(h^t)\sim\psi(h^t)}\sum_{t=0}^{\infty}\beta^t g(x_t, a_t(h^t), s_t)$$

$$=\sum_{a\in\mathcal{A}}\psi_0(a)\left[g(x_0, a, s_0) + \beta\mathbb{E}_{s_0}\mathbb{E}_{h^1=(s_0,a,s_1)}^{a(h^t)\sim\psi(h^t)}\sum_{t=1}^{\infty}\beta^{t-1}g(x_t, a_t, s_t)\right].$$

We take $\psi^* = \psi_0$ and $\phi'^*(a, s_1) = \mathbb{E}_{h^1=(s_0,a,s_1)}^{a(h^t)\sim\psi(h^t)}\sum_{t=1}^{\infty}\beta^{t-1}g(x_t, a_t, s_t)$. Hence (5.9) is satisfied.

It is straightforward to verify that, for any λ^* be a solution to (3.5) with $\gamma = 0$, and $\mu^* = 0$,

we have

$$\sum_{a\in\mathcal{A}}\psi^*(a)$$

$$\left[(r(x,a,s)+\lambda^*(g(x,a,s)-\bar{g})+\mu^*(g(x,a,s)-\phi))+\beta\mathbb{E}_s W(\lambda^*+\mu^*,x',s',\phi'^*(a,s'))\right]$$

$$\geq\mathbb{E}_{s_0}^{(a(h^t)\sim\psi(h^t))}\left[\sum_{t=0}^{\infty}\beta^t r(x_t,a_t,s_t)+\lambda^*(-\bar{g}+\sum_{t=0}^{\infty}\beta^t g(x_t,a_t,s_t))\right]$$

$$=\mathbb{E}_{s_0}^{(a(h^t)\sim\psi(h^t))}\left[\sum_{t=0}^{\infty}\beta^t r(x_t,a_t,s_t)\right]+\lambda^*\left(\sum_{a\in\mathcal{A}}\psi^*(a)\left(g(x,a,s)+\beta\mathbb{E}_s\phi'^*(a,s')-\bar{g}\right)\right)$$

$$\geq\mathbb{E}_{s_0}^{(a(h^t)\sim\psi(h^t))}\left[\sum_{t=0}^{\infty}\beta^t r(x_t,a_t,s_t)\right]=W(0,x,s,\phi),$$

(the last equality holds because of the relationship between $W(\gamma,x,s,\phi)$ and $W(0,x,s,\phi)$) with equality only if (5.10) is satisfied. Furthermore, ψ^*,ϕ'^* should be a solution to the sup problem given $\lambda^*,\mu^*=0$, and $\psi^*(h^t|s_0,a,s')(t\geq 1)$ is a solution to $W(\lambda^*+\mu^*,x',s',\phi'^*(a,s'))$, hence (5.8) shoule be satisfied according to lemma 5.1. Therefore $(\lambda^*,0,\psi^*,\phi'^*)$ is a solution to (5.7) satisfying (5.8), (5.9) and (5.10), and we finish the proof.

Step 2. Relation. Suppose that $(\lambda^*,\mu^*,\psi^*,\phi'^*)$ is a solution to (5.7) satisfying the (5.8),(5.9) and (5.10). Since $\phi'^*(a,s')\in\partial D(\lambda^*+\mu^*,\zeta(x,a,s),s')$, there exists $\psi^*(a,s')(\cdot)$ which is a solution to $W(\lambda^*+\gamma^*,x',s',\phi'^*(a,s'))$ and satisfies

$$\mathbb{E}_{s_0=s'}^{a(h^t)\sim\psi^*(a,s')(h^t)}\sum_{t=0}^{\infty}\beta^t g(x_t,a_t,s_t)=\phi'^*(a,s').$$

We define $\bar{\psi}^*(h^t)$ s.t. $\bar{\psi}^*(s_0)=\psi^*$ and $\bar{\psi}^*(s_0,a,s',\cdot)=\psi^*(a,s')(\cdot)$. It is then straightforward to verify that

$$W(0,x,s,\phi)=\sum_{a\in\mathcal{A}}\psi^*(a)$$

$$\left[(r(x,a,s)+\lambda^*(g(x,a,s)-\bar{g})+\mu^*(g(x,a,s)-\phi))+\beta\mathbb{E}_s W(\lambda^*+\mu^*,x',s',\phi'^*(a,s'))\right]$$

$$=\mathbb{E}_{s_0}^{(a(h^t)\sim\bar{\psi}^*(h^t))}\left[\sum_{t=0}^{\infty}\beta^t r(x_t,a_t,s_t)+\lambda^*(-\bar{g}+\sum_{t=0}^{\infty}\beta^t g(x_t,a_t,s_t))+\mu^*(-\phi+\sum_{t=0}^{\infty}\beta^t g(x_t,a_t,s_t))\right]$$

$$=\mathbb{E}_{s_0}^{(a(h^t)\sim\bar{\psi}^*(h^t))}\left[\sum_{t=0}^{\infty}\beta^t r(x_t,a_t,s_t)\right]+\lambda^*\left(\sum_{a\in\mathcal{A}}\psi^*(a)\left(g(x,a,s)+\beta\mathbb{E}_s\phi'^*(a,s')-\bar{g}\right)\right)$$

$$+\mu^*\left(\sum_{a\in\mathcal{A}}\psi^*(a)(g(x,a,s)+\beta\mathbb{E}_s\phi'^*(a,s')-\phi)\right)$$

$$=\mathbb{E}_{s_0}^{(a(h^t)\sim\bar{\psi}^*(h^t))}\left[\sum_{t=0}^{\infty}\beta^t r(x_t,a_t,s_t)\right],$$

due to the complementary conditions (5.9) and (5.10). And

$$\sum_{a \in \mathcal{A}} \psi^*(a)$$

$$\left[\left(r(x,a,s) + \gamma g(x,a,s) + \tilde{\lambda}^*(g(x,a,s) - \bar{g}) + \tilde{\mu}^*(g(x,a,s) - \phi) \right) + \beta \mathbb{E}_s W(\gamma + \tilde{\lambda}^* + \tilde{\mu}^*, x', s', \phi'^*(a,s')) \right]$$

$$\geq \mathbb{E}_{s_0}^{(a(h^t) \sim \bar{\psi}^*(h^t))} \left[\sum_{t=0}^{\infty} \beta^t r(x_t, a_t, s_t) + \sum_{t=0}^{\infty} \gamma g(x_t, a_t, s_t) + \right.$$

$$\left. + \tilde{\lambda}^*(-\bar{g} + \sum_{t=0}^{\infty} \beta^t g(x_t, a_t, s_t)) + \tilde{\mu}^*(-\phi + \sum_{t=0}^{\infty} \beta^t g(x_t, a_t, s_t)) \right]$$

$$= \mathbb{E}_{s_0}^{(a(h^t) \sim \bar{\psi}^*(h^t))} \left[\sum_{t=0}^{\infty} \beta^t r(x_t, a_t, s_t) \right] + \gamma \left(\sum_{a \in \mathcal{A}} \psi^*(a) \left(g(x,a,s) + \beta \mathbb{E}_s \phi'^*(a,s') \right) \right)$$

$$+ \tilde{\lambda}^* \left(\sum_{a \in \mathcal{A}} \psi^*(a) \left(g(x,a,s) + \beta \mathbb{E}_s \phi'^*(a,s') - \bar{g} \right) \right) + \tilde{\mu}^* \left(\sum_{a \in \mathcal{A}} \psi^*(a) \left(g(x,a,s) + \beta \mathbb{E}_s \phi'^*(a,s') - \phi \right) \right)$$

$$\geq \mathbb{E}_{s_0}^{(a(h^t) \sim \bar{\psi}^*(h^t))} \left[\sum_{t=0}^{\infty} \beta^t r(x_t, a_t, s_t) \right] + \gamma \phi = W(0, x, s, \phi) + \gamma \phi = W(\gamma, x, s, \phi),$$

with equality only if (5.5) and (5.6) are satisfied. Furthermore, ψ^*, ϕ'^* is a solution to the sup problem given $(\tilde{\lambda}^*, \tilde{\mu}^*)$ which is a solution to (5.3), and $\psi^*(a, s')(\cdot)$ is a solution to $W(\gamma + \lambda^* + \mu^*, x', s', \phi'^*(a, s'))$, hence Lemma 5.1 implies (5.4). Therefore $(\tilde{\lambda}^*, \tilde{\mu}^*, \psi^*, \phi'^*)$ is a solution to (5.3) satisfying (5.4), (5.5) and (5.6). $\qquad \square$

Proof of Theorem 5.6

- **0.** The convergence of Lagrangian multipliers follows from Proposition 2.7 in Nedic and Bertsekas (2001).

- **1.** Since the convergence of the Lagrangian multipliers, there exists $\bar{N}_1 > 0$, s.t. when $k \geq \bar{N}_1$,
$$\|(\lambda^k, \mu^k) - (\lambda^*, \mu^*)\| \leq \epsilon.$$

- **2.** According to the upper semi-continuity of the sub-gradient, there eixsts $\delta > 0$, s.t. when $a^k = a$ and
$$\|(\lambda^k, \mu^k) - (\lambda^*, \mu^*)\| \leq \delta,$$
we have
$$\text{dist}(\phi^k(s'), \partial D(\lambda^* + \mu^*, x', s')) \leq \frac{\epsilon}{2}.$$
According to the convergence of (λ^k, μ^k), there exists $\bar{M}_1 > 0$, s.t. when $k \geq \bar{M}_1$,

$$\|(\lambda^k, \mu^k) - (\lambda^*, \mu^*)\| \leq \delta.$$

Therefore, when $N > \bar{M}_1$, we have[16]

$$\text{dist}(\phi'^N(a, s'), \partial D(\lambda^* + \mu^*, x', s'))$$

$$= \text{dist}(\frac{\sum_{k=1}^{N} 1_{a^k=a} \sigma^k \phi^k(s')}{\sum_{k=1}^{N} 1_{a^k=a} \sigma^k}, \partial D(\lambda^* + \mu^*, x', s'))$$

$$\leq \sum_{k=1}^{\bar{M}_1} 1_{a^k=a} \tilde{\sigma}^k \text{dist}(\phi^k(s'), \partial D(\lambda^* + \mu^*, x', s')) + \sum_{k=\bar{M}_1+1}^{N} 1_{a^k=a} \tilde{\sigma}^k \text{dist}(\phi^k(s'), \partial D(\lambda^* + \mu^*, x', s'))$$

$$\leq \sum_{k=1}^{\bar{M}_1} 1_{a^k=a} \tilde{\sigma}^k \text{dist}(\phi^k(s'), \partial D(\lambda^* + \mu^*, x', s')) + \sum_{k=\bar{M}_1+1}^{N} 1_{a^k=a} \tilde{\sigma}^k \frac{\epsilon}{2}$$

$$\leq \sum_{k=1}^{\bar{M}_1} 1_{a^k=a} \tilde{\sigma}^k \text{dist}(\phi^k(s'), \partial D(\lambda^* + \mu^*, x', s')) + \frac{\epsilon}{2},$$

where

$$\tilde{\sigma}^k = \frac{\sigma^k}{\sum_{i=1}^{N} 1_{a^i=a} \sigma^i}.$$

Since

$$\psi^N(a) \tilde{\sigma}^k = \frac{\sum_{i=1}^{N} 1_{a^i=a} \sigma^i}{\sum_{i=1}^{N} \sigma^i} \cdot \frac{\sigma^k}{\sum_{i=1}^{N} 1_{a^i=a} \sigma^i} = \frac{\sigma^k}{\sum_{i=1}^{N} \sigma^i},$$

we have

$$\psi^N(a) \text{dist}(\phi'^N(a, s'), \partial D(\lambda^* + \mu^*, x', s'))$$

$$\leq \sum_{k=1}^{\bar{M}_1} \frac{\sigma^k}{\sum_{i=1}^{N} \sigma^i} \text{dist}(\phi^k(s'), \partial D(\lambda^* + \mu^*, x', s')) + \frac{\epsilon}{2}. \tag{A.41}$$

Since $\sum_{i=1}^{N} \sigma^i \to \infty$, there exits $\bar{N}_2 > \bar{M}_1$, s.t. when $N > \bar{N}_2$, the RHS of (A.41) $\leq \epsilon$.

- **3.** According to the updating rule for μ^k, we have

$$\mu^{k+1} = \max\{\mu^k - \sigma^k \left(g(x, a^k, s) - \phi + \beta \mathbb{E}_s \phi^k(s') \right), 0\}$$

$$\geq \mu^k - \sigma^k \left(g(x, a^k, s) - \phi + \beta \mathbb{E}_s \phi^k(s') \right). \tag{A.42}$$

Adding equations (A.42) for $k = 1, \cdots, N$, we have

$$\mu^{N+1} \geq \mu^1 - \sum_{k=M}^{N} \sigma^k \left(g(x, a^k, s) - \phi + \beta \mathbb{E}_s \phi^k(s') \right),$$

---

[16]The first inequality uses the fact that $\partial D(\lambda^* + \mu^*, x', s')$ is a convex set.

implying that

$$
\sum_a \psi^N(a)(g(x,a,s) + \beta \mathbb{E}_s \phi'^N(a,s')) - \phi
$$

$$
= \frac{\sum_a \left(\sum_{1 \le k \le N,\, a^k = a} \sigma^k \right) \left(g(x,a,s) - \phi + \beta \mathbb{E}_s \frac{\sum_{1 \le k \le N,\, a^k = a} \sigma^k \phi^k(s')}{\sum_{1 \le k \le N,\, a^k = a} \sigma^k} \right)}{\sum_{k=1}^N \sigma^k}
$$

$$
= \frac{\sum_a \left[\left(\sum_{1 \le k \le N,\, a^k = a} \sigma^k \right) (g(x,a,s) - \phi) + \beta \mathbb{E}_s \sum_{1 \le k \le N,\, a^k = a} \sigma^k \phi^k(s') \right]}{\sum_{k=1}^N \sigma^k}
$$

$$
= \frac{\sum_{k=1}^N \sigma^k \left(g(x,a^k,s) - \phi + \beta \mathbb{E}_s \phi^k(s') \right)}{\sum_{k=1}^N \sigma^k}
$$

$$
\ge \frac{\mu^1 - \mu^{N+1}}{\sum_{k=1}^N \sigma^k},
$$

where the right hand side tends to 0 as $N \to \infty$. Therefore, there exists $\bar{N}_3 > 0$, s.t. when $N > \bar{N}_3$, **(3.)** holds.

- **4.** Similar as **(3.)**, there exists $\bar{N}_4 > 0$, s.t. **(4.)** holds.

- **5.** It suffices to consider the case $\mu^* > 0$. Since $\mu^k \to \mu^*$, there exists $\bar{M}_2$, such that when $k \ge \bar{M}_2$, $\mu^k > 0$, and (A.42) holds as an equality. Therefore,

$$
\| \sum_a \psi^N(a)(g(x,a,s) + \beta \mathbb{E}_s \phi'^N(a,s')) - \phi \|
$$

$$
= \| \frac{\sum_{k=1}^N \sigma^k \left(g(x,a^k,s) - \phi + \beta \mathbb{E}_s \phi^k(s') \right)}{\sum_{k=1}^N \sigma^k} \|
$$

$$
= \| \frac{\sum_{k=1}^{\bar{M}_2} \sigma^k \left(g(x,a^k,s) - \phi + \beta \mathbb{E}_s \phi^k(s') \right) + \sum_{k=\bar{M}_2+1}^N \sigma^k \left(g(x,a^k,s) - \phi + \beta \mathbb{E}_s \phi^k(s') \right)}{\sum_{k=}^N \sigma^k} \|
$$

$$
\le \frac{\sum_{k=1}^{\bar{M}_2} \sigma^k}{\sum_{k=1}^N \sigma^k} \| g(x,a,s) - \phi + \beta \mathbb{E}_s \phi^k(s') \|_\infty + | \frac{\mu^{\bar{M}_2+1} - \mu^{N+1}}{\sum_{k=1}^N \sigma^k} | \to 0,
$$

implying that

$$
|\langle \mu^*, \sum_a \psi^N(a)(g(x,a,s) + \beta \mathbb{E}_s \phi'^N(a,s')) - \phi \rangle| \to 0, \text{ as } N \to \infty.
$$

Hence there exists $\bar{N}_5 > \bar{M}_2$, such that when $N \ge \bar{N}_5$, we have

$$
|\langle \mu^*, \sum_a \psi^N(a)(g(x,a,s) + \beta \mathbb{E}_s \phi'^N(a,s')) - \phi \rangle| < \epsilon.
$$

- **6.** Similar as **(5.)**, there exists $\bar{N}_6 > 0$, s.t. **(6.)** holds. Therefore, $\bar{N}$ can be taken as $\bar{N} = \max_{1 \le i \le 6} \{\bar{N}_i\}$. $\qquad \square$

Proof of Lemma 6.1

Since $c_s = \ell_s - g_s, \quad s \in \{1, \cdots, S\}$, we have

$$b = c_s - (1 - \tau_s)\ell_s = c_s + \frac{v'(\ell_s)}{u'(c_s)}\ell_s$$

$$= (\ell_s - g_s)(1 - \frac{\ell_s}{2\sqrt{1 - \ell_s}}) = f(\ell_s; g_s),$$

which is exactly (6.7). In period 0, since $c_0 = \ell_0 - g_0$, we have

$$b_{-1} = c_0 - (1 - \tau_0)\ell_0 + qb = c_0 + \frac{v'(\ell_0)}{u'(c_0)}\ell_0 + \beta\frac{p_1 u'(c_1) + p_2 u'(c_2)}{u'(c_0)}b$$

$$= (\ell_0 - g_0)(1 - \frac{\ell_0}{2\sqrt{1 - \ell_0}} + \beta p \frac{b}{\ell_1 - g_1} + \beta(1 - p)\frac{b}{\ell_2 - g_2})$$

$$= (\ell_0 - g_0)(1 - \frac{\ell_0}{2\sqrt{1 - \ell_0}} + \beta p \left(1 - \frac{\ell_1}{2\sqrt{1 - \ell_1}}\right) + \beta(1 - p)\left(1 - \frac{\ell_2}{2\sqrt{1 - \ell_2}}\right))$$

$$= f(\ell_0 - g_0; g_0) + (\ell_0 - g_0)(\underbrace{\beta p h(\ell_1) + \beta(1 - p)h(\ell_2)}_{w})).$$

Now we show that $2 \Rightarrow 1$. If $\ell_s > g_s$ satisfies (6.7) and (6.8) we construct

$$c_s = \ell_s - g_s, \ \tau_s = 1 + \frac{v'(\ell_s)}{u'(\ell_s - g_s)} < 1,$$

$$q = \frac{\beta p u'(\ell_1 - g_1) + \beta(1 - p)u'(\ell_2 - g_2)}{u'(\ell_0 - g_0)} > 0,$$

and it is easy to see that $\{(c_s, \ell_s, \tau_s)_{s=0}^S, q, b\}$ satisfies (6.2) to (6.5). $\qquad\square$

B Appendix: Preliminary Mathematical Results (on-line only)

B.1 Convex Conjugate

Definition B.1 *Let X be a Banach space, $f : X \to \mathbb{R} \cup \{+\infty, -\infty\}$ be an extended real-valued functional. If $f(x) \neq -\infty, \forall x \in X$, and f is not identically $+\infty$, then we call f a **proper** functional. The **effective domain** of f, denoted $dom(f)$, is defined by*

$$dom(f) = \{x \in X | f(x) < +\infty\}.$$

Definition B.2 *Let X be a Banach space, $f : X \to \mathbb{R} \cup \{+\infty, -\infty\}$ be an extended real-valued functional. The **epigraph** of f, denoted $epi(f)$, is defined by*

$$epi(f) := \{(x, r) \in X \times \mathbb{R} | f(x) \leq r\}.$$

Proposition B.1 *Let X be a Banach space, $f : X \to \mathbb{R} \cup \{+\infty, -\infty\}$ be an extended real-valued functional. Then*

$$f \text{ is convex} \Leftrightarrow epi(f) \text{ is convex in } X \times \mathbb{R}.$$

Proposition B.2 *Let X be a Banach space, $f : X \to \mathbb{R} \cup \{+\infty, -\infty\}$ be an extended real-valued functional. Then*

$$f \text{ is } \textbf{lower semicontinuous(l.s.c.)} \Leftrightarrow epi(f) \text{ is closed in } X \times \mathbb{R}.$$

Lemma B.1 *(Ascoli Theorem)*[17] *Let X be a Banach space, and $E \subset X$ be a closed convex set. Then for any $x_0 \notin E$, there exists $x^* \in X^*$ and $\gamma \in \mathbb{R}$, such that*

$$\langle x^*, x \rangle < \gamma < \langle x^*, x_0 \rangle, \quad \forall x \in E.$$

B.2   Optimization Problem, Lagrangian Dual

Definition B.3 *Let X, Ω, Y, f, g be defined as in Definition 4.1. We define the **perturbation functional** $v : Y \to \mathbb{R} \cup \{+\infty, -\infty\}$ by letting $v(y)$ be the optimal value of the following **perturbed problem***

$$
\begin{aligned}
\inf_{x \in \Omega} &\ f(x), \\
s.t. &\ g(x) \le y,
\end{aligned}
\tag{B.1}
$$

where $g(x) \le y$ means $g(x) - y \in -P$. In particular, $v(\theta_Y)$ is the optimal value of the optimization problem (4.1).

Definition B.4 *Let X, Ω, Y, f, g be defined as in Definition 4.1. The **Lagrangian function** $L : \Omega \times Y_+^* \to \mathbb{R} \cup \{+\infty\}$ to the optimization problem (4.1) is defined by*

$$L(x, y^*) = f(x) + \langle y^*, g(x) \rangle. \tag{B.2}$$

Here Y_+^* denotes the set

$$\{y^* \in Y^* | \langle y^*, y \rangle \ge 0, \ \forall y \ge \theta_Y \}.$$

Definition B.5 *Let X, Ω, Y, f, g be defined as in Definition 4.1. The **inf-sup problem**, or the **primal problem**, is defined by*

$$p := \inf_{x \in \Omega} \sup_{y^* \in Y_+^*} L(x, y^*). \tag{B.3}$$

*Similarly, the **sup-inf problem**, or the **dual problem**, is defined by*

$$d := \sup_{y^* \in Y_+^*} \inf_{x \in \Omega} L(x, y^*). \tag{B.4}$$

---

[17]This lemma can be deduced directly from the Hahn-Banach Theorem.

Remark B.1 *We can use Ω to absorb some constraints to consider the partial dual of an optimization problem. To be precise, if we have the optimization problem as following*

$$
\begin{aligned}
\inf_{x \in \Omega} \; & f(x), \\
s.t. \; & g(x) \leq \theta_Y, \\
& h(x) \leq \theta_Z.
\end{aligned}
\tag{B.5}
$$

Then we can define

$$
\tilde{\Omega} := \{x \in \Omega \mid h(x) \leq \theta_Z\} \subseteq \Omega \subseteq X,
$$

and rewrite the optimization problem (B.5) *as*

$$
\begin{aligned}
\inf_{x \in \tilde{\Omega}} \; & f(x), \\
s.t. \; & g(x) \leq \theta_Y.
\end{aligned}
\tag{B.6}
$$

and apply theorem 4.2 to problem (B.6).

Corollary B.1 *Assume that $X = \mathbb{R}^m$, $\Omega = X$, $Y = \mathbb{R}^n$, and f is proper and bounded from below, i.e. there exists $M \in \mathbb{R}$, s.t. $f(x) \geq M$, $\forall x \in X$. Then for any $\epsilon > 0$, there exists $y_\epsilon \in Y$, $\|y_\epsilon\| \leq \epsilon$, $\{p_i \geq 0\}_{i=1}^{n+2}$, and $\{x_i \in X\}_{i=1}^{n+2}$, s.t.*

$$
\sum_{i=1}^{n+2} p_i = 1, \quad \sum_{i=1}^{n+2} p_i g(x_i) - y_\epsilon \leq \theta_Y,
$$

and

$$
\sum_{i=1}^{n+2} p_i f(x_i) \leq d + \epsilon.
$$

Proof Since f is proper and bounded from below, we know that the perturbation function v is also proper and bounded from below, therefore also satisfies Assumption **??**. According to Theorem 4.1, we have

$$
\mathrm{epi}(v^{**}) = \mathrm{cl} \; \mathrm{co} \; \mathrm{epi}(v).
$$

Note that $(\theta_Y, d) \in \mathrm{epi}(v^{**})$ according to Theorem 4.2, hence there exists $(\bar{y}, \bar{v}) \in \mathrm{co} \; \mathrm{epi}(v) \subset Y \times \mathbb{R}$, such that $\|(\bar{y}, \bar{v}) - (\theta_Y, d)\|_{Y \times \mathbb{R}} < \frac{\epsilon}{2}$, implying that

$$
\|\bar{y}\|_Y < \frac{\epsilon}{2} < \epsilon, \quad \|\bar{v} - d\| < \frac{\epsilon}{2}.
$$

By Carathéodory's Theorem, since $(\bar{y}, \bar{v}) \in \mathrm{co} \; \mathrm{epi}(v) \subset \mathbb{R}^{n+1}$, there exists $\{p_i \geq 0\}_{i=1}^{n+2}$, $\{(y_i, v_i) \in \mathrm{epi}(v)\}_{i=1}^{n+2}$, s.t.

$$
\sum_{i=1}^{n+2} p_i = 1, \quad \sum_{i=1}^{n+2} p_i (y_i, v_i) = (\bar{y}, \bar{v}).
$$

By the definition of v, for any $i \in \{1, \cdots, n+2\}$, there exisits $x_i \in X$, such that

$$g(x_i) \le y_i, \quad f(x_i) \le v_i + \frac{\epsilon}{2}.$$

Therefore, we have

$$\sum_{i=1}^{n+2} p_i g_i(x_i) - \bar{y} \le \sum_{i=1}^{n+2} p_i y_i - \bar{y} = \theta_Y,$$

and

$$\sum_{i=1}^{n+2} p_i f(x_i) \le \sum_{i=1}^{n_2} p_i (v_i + \frac{\epsilon}{2}) = \bar{v} + \frac{\epsilon}{2} \le d + \epsilon.$$

Hence we finish the proof. $\qquad\square$

Corollary B.2 *Assume that X is a Banach space, $\Omega = X$, Y is a Banach space with a closed positive cone $P \subset Y$, and f is proper and bounded from below, i.e. there exists $M \in \mathbb{R}$, s.t. $f(x) \ge M$, $\forall x \in X$. Then for any $\epsilon > 0$, there exists $y_\epsilon \in Y$, $\|y_\epsilon\| \le \epsilon$, $N \in \mathbb{N}_+$, $\{p_i \ge 0\}_{i=1}^N$, and $\{x_i \in X\}_{i=1}^N$, s.t.*

$$\sum_{i=1}^N p_i = 1, \quad \sum_{i=1}^N p_i g(x_i) - y_\epsilon \le \theta_Y,$$

and

$$\sum_{i=1}^N p_i f(x_i) \le d + \epsilon.$$

Proof Similarly to the proof for Corollary B.1, we can still find $(\bar{y}, \bar{v}) \in \text{co epi}(v) \subset Y \times \mathbb{R}$, such that

$$\|\bar{y}\|_Y < \frac{\epsilon}{2} < \epsilon, \quad \|\bar{v} - d\| < \frac{\epsilon}{2}.$$

Note that co epi(v) can be expressed by

$$\text{co epi}(v) = \{\sum_{i=1}^n p_i(y_i, v_i) | n \in \mathbb{N}_+, \ p_i \ge 0, \ (y_i, v_i) \in \text{epi}(v), \ \sum_{i=1}^n p_i = 1\}.$$

Hence there exists $N \in \mathbb{N}_+$, $\{p_i \ge 0\}_{i=1}^N$, $\{(y_i, v_i) \in \text{epi}(v)\}_{i=1}^N$, s.t.

$$\sum_{i=1}^N p_i = 1, \quad \sum_{i=1}^N p_i(y_i, v_i) = (\bar{y}, \bar{v}).$$

The rest of the proof is the same as the proof for Corollary B.1, and the details are omitted here. $\qquad\square$

Corollary B.3 *Assume that X is a Banach space, $\Omega = X$, Y is a Banach space with a closed positive cone $P \subset Y$, and f is proper and bounded from below, i.e. there exists $M \in \mathbb{R}$, s.t. $f(x) \ge M$, $\forall x \in X$. Then for any $N \in \mathbb{N}_+$, $\{p_i \ge 0\}_{i=1}^N$, and $\{x_i \in X\}_{i=1}^N$, s.t.*

$$\sum_{i=1}^N p_i = 1, \quad \sum_{i=1}^N p_i g(x_i) \le \theta_Y,$$

we have

$$\sum_{i=1}^{N} p_i f(x_i) \geq d.$$

Proof For any $N \in \mathbb{N}_+$, s.t.

$$\sum_{i=1}^{N} p_i = 1, \quad \sum_{i=1}^{N} p_i g(x_i) \leq \theta_Y,$$

it is straightforward to verify that $(y_i, f(x_i)) \in \text{epi}(v)$ for any $i \in \{1, \cdots, N\}$, $y_i \geq g(x_i)$ by the definition of the perturbation functional v. Therefore,

$$(\sum_{i=1}^{N} p_i y_i, \sum_{i=1}^{N} p_i f(x_i)) \in \text{co epi}(v), \quad \forall y_i \geq g(x_i).$$

Since

$$\sum_{i=1}^{N} p_i g(x_i) \leq \theta_Y,$$

we have

$$(\theta_Y, \sum_{i=1}^{N} p_i f(x_i)) \in \text{co epi}(v) \subset \text{cl co epi} = \text{epi}(v^{**}),$$

implying that

$$\sum_{i=1}^{N} p_i f(x_i) \geq v^{**}(\theta_Y) = d,$$

according to Theorem 4.2. $\qquad\qquad\square$

C   Existence of Lagrange Multiplier

In this appendix we give a detailed review on the existence of Lagrange mulltilpiers in appropriate function spaces. Most of the material can be found in Dechert (1982) and Pavoni et al. (2018b). It is included here for completeness.

Definition C.1 *Let $f : \ell^\infty \to \mathbb{R}$ be a functional bounded from below, i.e. there exists $M \in \mathbb{R}$, s.t. $f(x) \geq M$, $\forall x \in \ell^\infty$; and $g : \ell^\infty \to \ell^\infty$ be an arbitrary functional. The ℓ^∞-**optimization problem**, is defined by*

$$\inf_{x \in \ell^\infty} f(x), \tag{C.1}$$
$$s.t.\ g(x) \leq 0,$$

where $g(x) \leq 0$ means $g_t(x) \leq 0$, $\forall t \in \mathbb{Z}^+$.

Lemma C.1 *(Yosida-Hewitt Decomposition Theorem, see Yosida and Hewitt (1952)(Theorem 1.23, Theorem 1.24, Theorem 2.3)) For any $\lambda \in \ell^{\infty,*}$(or $\ell_+^{\infty,*}$), there exists a unique $\lambda^1 \in \ell^1$(or ℓ_+^1), and a unique $\lambda^s \in \ell^s$(or ℓ_+^s), s.t. $\lambda = \lambda^1 + \lambda^s$. Here ℓ^s is the set of **purely finitely additive measures**.*

Remark C.1   *1. A well-known example of $\lambda \in \ell^{\infty,*} - \ell^1$ is constructed by Hahn-Banach Theorem. To be precise, we define the set $c \subset \ell^\infty$ as*

$$c = \{x \in \ell^\infty | \lim_{t \to \infty} x_t \text{ exists}\}.$$

Therefore the operator f_c defined by

$$f_c(x) = \lim_{t \to \infty} x_t, \quad \forall x \in c,$$

is a bounded linear operator on c. According to Hahn-Banach Theorem, this operator f_c can be extended to be a bounded linear operator on ℓ^∞, or equivalently, to be an element of $\ell^{\infty,*}$, denoted as f. It is straightforward to check that $f \notin \ell^1$.

2. *Another approach to construct an element belonging to $\ell^{\infty,*} - \ell^1$ is to use **ultrafilters**. To be precise, for any $A \subseteq \mathbb{Z}^+$, let e_A be the indicator function of A, i.e.,*

$$e_A(t) = \begin{cases} 1, & t \in A; \\ 0, & t \notin A. \end{cases}$$

We consider the set[18]

$$M := \{\lambda \in \ell^{\infty,*} | \langle \lambda, e_{\mathbb{Z}^+} \rangle = 1; \langle \lambda, e_A \rangle = 0 \text{ or } 1, \forall A \subset \mathbb{Z}^+\}.$$

We note that when $\lambda \in M$, such that $\langle \lambda, e_A \rangle = 0$ for any finite sets $A \subset \mathbb{Z}^+$, then $\lambda \in \ell^{\infty,*} - \ell^1$.

3. *We define $c_0 \subset c \subset \ell^\infty$ as*

$$c_0 = \{x \in \ell^\infty | \lim_{t \to \infty} x_t = 0\}.$$

Then $c_0^* = \ell^1$. Various asymptotic behaviors of $x \in \ell^\infty$ bring bounded linear functionals in $\ell^{\infty,*} - \ell^1$. This is why we need restrictions on asymptotic behaviors of g below for further discussions.

Definition C.2 *Let $g : \ell^\infty \to \ell^\infty$ be a functional. For any $u, v \in \ell^\infty$, let*

$$x^T(u, v) := \begin{cases} u_t, & t \leq T; \\ v_t, & t > T. \end{cases}$$

*We say the functional g is **Asymptotically Insensitive(AI)**, if*

$$\lim_{t \to \infty} \left[g_t(x^T(u, v)) - g_t(v) \right] = 0, \quad \forall u, v \in \ell^\infty, T \in \mathbb{Z}^+.$$

---

[18]As shown in Dechert (1982)(Page 289), M is indeed the set of extreme points of

$$G := \{\lambda \in \ell^{\infty,*} | \langle \lambda, e_{\mathbb{Z}^+} \rangle = 1\}.$$

.

*We say the funtional g is **Asymptotoically Non-Anticipatory(ANA)** if*

$$\lim_{T \to \infty} \left[g_t(x^T(u,v)) - g_t(u) \right] = 0, \quad \forall u, \, v \in \ell^\infty, \, t \in \mathbb{Z}^+.$$

Remark C.2   *1. These terminologies were first discussed in Dechert (1982) and were adapted for the dual approach to dynamic problems in Pavoni et al. (2018a). Here we use the definitions in Pavoni et al. (2018a).*

    2. Intuitively, g is AI means that for any $t \in \mathbb{Z}^+$, g_t is not quite related to x_r with $r \ll t$; and g is ANA means that for any $t \in \mathbb{Z}^+$, g_t is not quite related to x_r with $r \gg t$. A typical type of constraint in dynamic problems is

$$g_t^D(x) := \sum_{r=0}^\infty \beta^r h_{t+r}(x_{t+r}) \leq 0, \quad \forall t \in \mathbb{Z}^+,$$

where $0 < \beta < 1$ denotes the discount factor. It is straightforward to check that g^D is both AI and ANA under some mild regularity assumptions on h.

Lemma C.2 *Let $g : \ell^\infty \to \ell^\infty$ be AI. Then for any $\lambda^s \in \ell^s$, $u, \, v \in \ell^\infty$, $T \in \mathbb{Z}^+$, we have*

$$\langle \lambda^s, g(x^T(u,v)) \rangle = \langle \lambda^s, g(v) \rangle.$$

Proof See Pavoni et al. (2018b)(Page 11, Lemma C.1). $\square$

Lemma C.3 *Let $g : \ell^\infty \to \ell^\infty$ be ANA. Given $u, \, v \in \ell^\infty$. Moreover, assume that $g(x^T(u,v))$ is uniformly bounded in ℓ^∞ with respect to $T \in \mathbb{Z}^+$. Then for any $\lambda^1 \in \ell^1$, we have*

$$\lim_{T \to \infty} \langle \lambda^1, g(x^T(u,v)) \rangle = \langle \lambda^1, g(u) \rangle.$$

Proof See Pavoni et al. (2018b)(Page 11, Lemma C.2). $\square$

Theorem C.1 *Let $f : \ell^\infty \to \mathbb{R}$ satisfy*

$$\lim_{T \to \infty} f(x^T(u,v)) = f(u),$$

for any $u, v \in \ell^\infty$. Let $g : \ell^\infty \to \ell^\infty$ be AI and ANA, and that there exists $u_0 \in \ell^\infty$, such that $\sup \lim_{t \to \infty} g_t(u_0) \leq 0^{19}$. Assume also that, for any $u, \, v \in \ell^\infty$, $g(x^T(u,v))$ is uniformly bounded in ℓ^∞ with respect to $T \in \mathbb{Z}^+$. Then the dual problem of (C.1) satisfies

$$d = \sup_{\lambda \in \ell_+^{\infty,*}} \inf_{x \in \ell^\infty} f(x) + \langle \lambda, g(x) \rangle = \sup_{\lambda \in \ell_+^1} \inf_{x \in \ell^\infty} f(x) + \langle \lambda, g(x) \rangle. \tag{C.2}$$

---

[19]This assumption is weaker than the assumption that there exists at least one feasible point to the problem (C.1). Indeed, if there exists a feasible point v_0 s.t. $g(v_0) \leq 0$, then obviously v_0 satisfies $\sup \lim_{t \to \infty} g_t(u_0) \leq 0$.

Proof • For step 1, we aim to show that

$$\inf_{x \in \ell^\infty} f(x) + \langle \lambda^1 + \lambda^s, g(x) \rangle \leq \inf_{x \in \ell^\infty} f(x) + \langle \lambda^1, g(x) \rangle, \quad \forall \lambda^1 \in \ell^1_+, \lambda^s \in \ell^s_+. \quad \text{(C.3)}$$

Given $\lambda^1 \in \ell^1_+$, $\lambda^s \in \ell^s_+$. For any $u \in \ell^\infty$, according to Lemma C.2, we know that

$$\langle \lambda^s, g(x^T(u, u_0)) \rangle = \langle \lambda^s, g(u_0) \rangle \leq 0.$$

Therefore,

$$\begin{aligned}
&\inf_{x \in \ell^\infty} f(x) + \langle \lambda^1 + \lambda^s, g(x) \rangle \\
&\leq f(x^T(u, u_0)) + \langle \lambda^1 + \lambda^s, g(x^T(u, u_0)) \rangle \\
&\leq f(x^T(u, u_0)) + \langle \lambda^1, g(x^T(u, u_0)) \rangle, \quad \forall u \in \ell^\infty, T \in \mathbb{Z}^+.
\end{aligned} \quad \text{(C.4)}$$

According to Lemma C.3, we know that

$$\lim_{T \to \infty} \langle \lambda^1, g(x^T(u, u_0)) \rangle = \langle \lambda^1, g(u) \rangle.$$

Therefore, we take the limit $T \to \infty$ in (C.4) and obtain

$$\begin{aligned}
&\inf_{x \in \ell^\infty} f(x) + \langle \lambda^1 + \lambda^s, g(x) \rangle \\
&\leq \lim_{T \to \infty} f(x^T(u, u_0)) + \langle \lambda^1, g(x^T(u, u_0)) \rangle \\
&= f(u) + \langle \lambda^1, g(u) \rangle, \quad \forall u \in \ell^\infty.
\end{aligned} \quad \text{(C.5)}$$

Since (λ^1, λ^s) in (C.5) is arbitrarily chosen from (ℓ^1_+, ℓ^s_+), we can conclude (C.3) from (C.5).

• For step 2, we utilize Yosida-Hewitt Decomposition to finish the proof. According to Lemma C.1, any $\lambda \in \ell^{\infty,*}_+$ can be decomposed as $\lambda = \lambda^1 + \lambda^s$, where $\lambda^1 \in \ell^1_+$ and $\lambda^s \in \ell^s_+$. It is then straightforward to conclude (C.2) from (C.3). $\qquad \square$

Remark C.3 *Under the same assumptions as Theorem C.1, but with the assumption of the existence of $u_0 \in \ell^\infty$ s.t. $\sup \lim_{t \to \infty} g_t(u_0) \leq 0$ replaced by the **Slater's condition**, it is shown in Pavoni et al. (2018b)(Page 11, Theorem C.2) that if $(x^*, \lambda^*) \in \ell^\infty \times \ell^{\infty,*}$ is a saddle point of $L(x, \lambda) = f(x) + \langle \lambda, g(x) \rangle$, then $(x^*, \lambda^*) \in \ell^\infty \times \ell^1$. The properties of the saddle points for the Lagrangian are studied in Dechert (1982) and Pavoni et al. (2018a), while Theorem C.1 addresses the dual problem.*

In practical applications, it is more common to consider optimization problems with a bounded feasible set $\mathcal{A} \subset \ell^\infty$. The requirement for f in Theorem C.1 and the ANA property for g are easily satisfied **uniformly** with respect to $u \in \mathcal{A}$, $v \in \ell^\infty$, s.t. $x^T(u, v) \in \mathcal{A}$ based on the existence of a discount factor $0 < \beta < 1$ over periods. However, the AI property for g is mainly satisfied by problems with only forward constraints and backward constraints that only

64

affect a few periods. Therefore, for problems with a bounded feasible set $\mathcal{A} \subset \ell^\infty$, a discount factor $0 < \beta < 1$ across periods, forward-looking constraints, and general **backward-looking constraints**, different assumptions should be made.

Let $\mathcal{A} \subset \ell^\infty$. For any $u \in \mathcal{A}$, we define

$$\mathcal{B}_T(u) = \{v \in \ell^\infty | x^T(u, v) \in \mathcal{A}\}.$$

In this section, we consider a slightly different problem to (C.1) as following

$$\inf_{x \in \mathcal{A}} f(x),$$
$$\text{s.t. } g(x) \leq 0. \tag{C.6}$$

Definition C.3 *We say the functional $g : \mathcal{A} \to \ell^\infty$ is **Uniformly Asymptotically Non-Anticipatory(UANA)**, if for any $t \in \mathbb{Z}^+$ and $u \in \mathcal{A}$,*

$$\lim_{T \to \infty} \sup_{v \in \mathcal{B}_T(u)} |g_t(x^T(u, v)) - g_t(u)| = 0$$

Lemma C.4 *Let $g : \mathcal{A} \to \ell^\infty$ be UANA. Given $u \in \mathcal{A}$, assume additionaly that $g(x^T(u, v))$ is uniformly bounded in ℓ^∞ with respect to $T \in \mathbb{Z}^+$ and $v \in \mathcal{B}_T(u)$. Then for any $\lambda^1 \in \ell^1$, we have*

$$\lim_{T \to \infty} \sup_{v \in \mathcal{B}_T(u)} |\langle \lambda^1, g(x^T(u, v)) \rangle - \langle \lambda^1, g(u) \rangle| = 0$$

Proof Since $g(x^T(u, v))$ is uniformly bounded in ℓ^∞, there exists $M > 0$, s.t.

$$\|g(x^T(u, v)) - g(u)\|_{\ell^\infty} < M, \quad \forall T \in \mathbb{Z}^+, \, v \in \mathcal{B}_T(u).$$

Since $\lambda^1 \in \ell^1$, for any $\epsilon > 0$, there exists $\bar{t} > 0$, such that

$$\sum_{t=\bar{t}+1}^\infty |\lambda_t^1| < \frac{\epsilon}{2M}.$$

Since g is UANA, there exists $\bar{T} > 0$, such that

$$|g_t(x^T(u, v)) - g_t(u)| < \frac{\epsilon}{2\|\lambda^1\|_{\ell^1}}, \quad \forall T \geq \bar{T}, \, v \in \mathcal{B}_T(u), \, t \leq \bar{t}.$$

Therefore,

$$|\langle \lambda^1, g(x^T(u, v)) - g(u) \rangle|$$
$$\leq \sum_{t=1}^{\bar{t}} \|\lambda^1\|_{\ell^1} |g_t(x^T(u, v)) - g_t(u)| + \sum_{t=\bar{t}+1}^\infty |\lambda_t^1| \|g(x^T(u, v)) - g(u)\|_{\ell^\infty}$$
$$\leq \|\lambda^1\|_{\ell^1} \frac{\epsilon}{2\|\lambda^1\|_\ell^1} + \frac{\epsilon}{2M} \cdot M = \epsilon, \quad \forall T \geq \bar{T}, \, v \in \mathcal{B}_T(u).$$

65

Since ϵ is chosen arbitrarily, we finish the proof. $\qquad\square$

Assumption C.1 *We assume that for any $u \in \mathcal{A}$ and $\bar{T} \in \mathbb{Z}^+$, there exists $T \geq \bar{T}$ and $v \in \mathcal{B}_T(u)$, s.t.*

$$\limsup_{t \to \infty} g_t(x^T(u,v)) \leq 0.$$

Theorem C.2 *Let $f : \mathcal{A} \to \mathbb{R}$ satisfy*

$$\lim_{T \to \infty} \sup_{v \in \mathcal{B}_T(u)} |f(x^T(u,v)) - f(u)| = 0. \tag{C.7}$$

Let $g : \mathcal{A} \to \ell^\infty$ be UANA, and satisfy Assumption C.1. Assume additionaly that, for any $u \in \mathcal{A}$, $g(x^T(u,v))$ is uniformly bounded in ℓ^∞ with respect to $T \in \mathbb{Z}^+$ and $v \in \mathcal{B}_T(u)$. Then the dual problem of (C.1) satisfies

$$d = \sup_{\lambda \in \ell^{\infty,*}_+} \inf_{x \in \mathcal{A}} f(x) + \langle \lambda, g(x) \rangle = \sup_{\lambda \in \ell^1_+} \inf_{x \in \mathcal{A}} f(x) + \langle \lambda, g(x) \rangle. \tag{C.8}$$

Proof The spirit of the proof is indeed the same as the proof for Theorem C.1. We again divided it into two steps.

- For step 1, we aim to show that

$$\inf_{x \in \mathcal{A}} f(x) + \langle \lambda^1 + \lambda^s, g(x) \rangle \leq \inf_{x \in \mathcal{A}} f(x) + \langle \lambda^1, g(x) \rangle, \quad \forall \lambda^1 \in \ell^1_+, \ \lambda^s \in \ell^s_+. \tag{C.9}$$

  Given $\lambda^1 \in \ell^1_+$, $\lambda^s \in \ell^s_+$. For any $u \in \mathcal{A}$ and $\bar{T} > 0$, according Assumption C.1, we know that there exists $T \geq \bar{T}$ and $v(u,\bar{T}) \in \mathcal{B}_T(u)$, s.t.

$$\langle \lambda^s, g(x^T(u, v(u,\bar{T}))) \rangle \leq 0.$$

  Therefore,

$$\begin{aligned}
&\inf_{x \in \mathcal{A}} f(x) + \langle \lambda^1 + \lambda^s, g(x) \rangle \\
&\leq f(x^T(u, v(u,\bar{T}))) + \langle \lambda^1 + \lambda^s, g(x^T(u, v(u,\bar{T}))) \rangle \\
&\leq f(x^T(u, v(u,\bar{T}))) + \langle \lambda^1, g(x^T(u, v(u,\bar{T}))) \rangle, \quad \forall u \in \mathcal{A}, \ \bar{T} \in \mathbb{Z}^+.
\end{aligned} \tag{C.10}$$

  According to Lemma C.4, and the assumption that

$$\lim_{T \to \infty} \sup_{v \in \mathcal{B}_T(u)} |f(x^T(u,v)) - f(u)| = 0,$$

  we take the limit $T \to \infty$ in (C.10) and obtain

$$\begin{aligned}
&\inf_{x \in \mathcal{A}} f(x) + \langle \lambda^1 + \lambda^s, g(x) \rangle \\
&\leq f(u) + \langle \lambda^1, g(u) \rangle, \quad \forall u \in \mathcal{A}.
\end{aligned} \tag{C.11}$$

66

Since (λ^1, λ^s) in (C.11) is arbitrarily chosen from (ℓ_+^1, ℓ_+^s), we can conclude (C.9) from (C.11).

- For step 2, we utilize Yosida-Hewitt Decomposition to finish the proof. According to Lemma C.1, any $\lambda \in \ell_+^{\infty,*}$ can be decomposed as $\lambda = \lambda^1 + \lambda^s$, where $\lambda^1 \in \ell_+^1$ and $\lambda^s \in \ell_+^s$. It is then straightforward to conclude (C.8) from (C.9). $\qquad\square$

The following theorem establishes the existence for the dual problem.

Theorem C.3 *Under Assumption 3.1, the dual Lagrangian problem of* (4.6) *can be formulated as*

$$
d = \inf_{(\lambda_t^i(h^t)) \in \Lambda} \sup_{(a(s^t)) \in \tilde{\mathcal{A}}^\infty(x_0)} L((a(s^t)), (\lambda^i(h^t)); (\gamma^i), x_0, s_0), \tag{C.12}
$$

where L is the Lagrangian functional defined in (3.3).

Proof The problem (4.6) can be reformulated as

$$
\max_{(a(s^t)) \in \tilde{\mathcal{A}}^\infty(x_0) \subset \mathcal{A}^\infty} f((a(s^t))),
$$
$$
\textbf{s.t. } g((a(s^t))) \geq 0, \tag{C.13}
$$

where f and g are defined in (4.7) and (4.8). It then suffices to verify all the conditions in Theorem C.2: Due to $0 < \beta < 1$, and the boundedness of r and g^i, it is direct to verify that f satisfies (C.7), g is UANA, and g is bounded in ℓ^∞. For any $u = (u_t(s^t)) \in \tilde{\mathcal{A}}^\infty(x_0)$, we denote x_t^u the state induced by u. For any $T \in \mathbb{N}, s^T \in \mathcal{S}^T$, according to the fifth assumption in Assumption 3.1, the problem with $(x_0' = x_T^u(s^T), s_0' = s_{T+1})$ has a feasible point. Therefore, there exists $v = (v_t(s^t))$, s.t. $v_t(s^T, s^{t-T})_{t \geq T} \in \tilde{A}^\infty(x_T^u(s^T))$ for any $s^T \in \mathcal{S}^T$, such that when considering $a^T(u, v) = ([a^T(u, v)]_t(s^t)) \in \tilde{\mathcal{A}}^\infty(x_0)$ defined as

$$
[a^T(u, v)]_t(s^t) = \begin{cases} u_t(s^t), & t \leq T; \\ v_t(s^t), & t > T, \end{cases}
$$

$(g(a^T(u, v)))_{t, h^t, i} \geq 0$ when $t \geq T$. Therefore, Theorem C.1 applies. $\qquad\square$

Note that in this formulation, the multipliers that correspond to the forward looking constraint starting that time t depend on the history up to time t but cannot be conditioned on the action taken at time t.

D   Relation to Bloise and Siconolfi (2022)

We consider a one-principle one-agent problem

$$V(x_0, s_0, \mu_0) = \sup_{\{a_t\}} \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t (u(x_t, a_t, s_t) + \mu_0 v(x_t, a_t, s_t))$$

$$\text{s.t. } x_{t+1} = \zeta(x_t, a_t, s_t), \quad p(x_t, a_t, s_t) \geq 0, \quad \forall t \geq 0, \, s^t \in \mathcal{S}^t$$

$$\mathbb{E}_t \sum_{n=0}^{\infty} \beta^n v(x_{t+n}, a_{t+n}, s_{t+n}) \geq \bar{v}, \quad \forall t \geq 0, \, s^t \in \mathcal{S}^t.$$

We define

$$W^1(\mathbb{E}_s u', x, a, s) = u(x, a, s) + \beta \mathbb{E}_s u', \, W^2(\mathbb{E}_s v', x, a, s) = v(x, a, s) + \beta \mathbb{E}_s v'.$$

Applying the Negishi's method in Bloise and Siconolfi (2022), we have the recursive formula as follows

$$J(\theta^1, \theta^2, x, s) = \sup_{a, u', v'} \theta^1 W^1(\mathbb{E}_s u', x, a, s) + \theta^2 W^2(\mathbb{E}_s v', x, a, s)),$$

$$\text{s.t. } p(x, a, s) \geq 0, x' = \zeta(x, a, s),$$

$$v(x, a, s) + \beta \mathbb{E}_s v' \geq \bar{v},$$

$$(u'(x', s'), v'(x', s')) \in (\mathcal{U}(x', s')) = \{\tilde{\theta}^1 u'(x', s') + \tilde{\theta}^2 v'(x', s') \leq J(\tilde{\theta}^1, \tilde{\theta}^2, x', s'), \forall \tilde{\theta}^1, \tilde{\theta}^2 \geq 0\}$$

It is straightforward to see that $\mathcal{U}(x', s')$ is closed and convex for any $x' \in \mathcal{X}$, $s' \in \mathcal{S}$. Furthermore, J is homogeneous of degree 1 w.r.t. (θ^1, θ^2), i.e. $J(t\theta^1, t\theta^2, x, s) = tJ(\theta^1, \theta^2, x, s)$ for all $t \geq 0$; and it is straightforward to see that J is convex w.r.t. (θ_1, θ_2). Fix $\theta^1 = 1$, We define $D(\theta^2, x, s) = J(1, \theta^2, x, s)$. Then D is convex w.r.t. θ^2. The Negishi's recursive formula for D is written

$$D(\theta^2, x, s) = \sup_{a, u', v'} W^1(\mathbb{E}_s u', x, a, s) + \theta^2 W^2(\mathbb{E}_s v', x, a, s)),$$

$$\text{s.t. } p(x, a, s) \geq 0, x' = \zeta(x, a, s),$$

$$v(x, a, s) + \beta \mathbb{E}_s v' \geq \bar{v},$$

$$(u'(x', s'), v'(x', s')) \in (\mathcal{U}(x', s')) = \{u'(x', s') + \tilde{\theta}^2 v'(x', s') \leq D(\tilde{\theta}^2, x', s'), \forall \tilde{\theta}^2 \geq 0\}$$
(D.1)

We define the set of all feasible (a, u', v') as $\mathcal{F}$. Problem (D.1) can be rewritten as

$$D(\theta^2, x, s) = \sup_{(a, u', v') \in \mathcal{F}} F(a, u, v') := u(x, a, s) + \theta^2 v(x, a, s) + \beta \mathbb{E}_s u' + \theta^2 \beta \mathbb{E}_s v' \qquad \text{(D.2)}$$

1. **We show that**

$$D(\theta^2, x, s) \le \sup_{a \in \tilde{\mathcal{A}}(x,s)} \inf_{\lambda \ge 0} u(x, a, s) + \theta^2(v(x, a, s) + \lambda(v(x, a, s) - \bar{v}) + \beta \mathbb{E}_s D(\theta^2 + \lambda, x, s).$$
(D.3)

Noticing that $\forall (a, u', v') \in \mathcal{F}$ and $\lambda \ge 0$, we have

$$\begin{aligned}
F(a, u', v') &= u(x, a, s) + \theta^2 v(x, a, s) + \beta \mathbb{E}_s u' + \theta^2 \beta \mathbb{E}_s v' \\
&\le u(x, a, s) + \theta^2 v(x, a, s) + \lambda(v(x, a, s) + \beta \mathbb{E}_s v' - \bar{v}) + \beta \mathbb{E}_s u' + \theta^2 \beta \mathbb{E}_s v' \\
&= u(x, a, s) + \theta^2(v(x, a, s) + \lambda(v(x, a, s) - \bar{v}) + \beta \mathbb{E}_s(u' + (\theta^2 + \lambda)v') \\
&\le u(x, a, s) + \theta^2(v(x, a, s) + \lambda(v(x, a, s) - \bar{v}) + \beta \mathbb{E}_s D(\theta^2 + \lambda, x', s').
\end{aligned}$$

Since λ is arbitrarily chosen, we have

$$F(a, u', v') \le \inf_{\lambda} u(x, a, s) + \theta^2(v(x, a, s) + \lambda(v(x, a, s) - \bar{v}) + \beta \mathbb{E}_s D(\theta^2 + \lambda, x, s),$$

for all $(a, u', v') \in \mathcal{F}$, implying that

$$D(\theta^2, x, s) = \sup_{(a,u',v') \in \mathcal{F}} F(a, u', v') \le \sup_{(a,u',v') \in \mathcal{F}} \inf_{\lambda} u(x, a, s) + \theta^2(v(x, a, s) + \lambda(v(x, a, s) - \bar{v}) + \beta \mathbb{E}_s D(\theta^2 + \lambda, x, s)$$

Therefore,

$$D(\theta^2, x, s) \le \sup_{a \in \tilde{\mathcal{A}}(x,s)} \inf_{\lambda \ge 0} u(x, a, s) + \theta^2(v(x, a, s) + \lambda(v(x, a, s) - \bar{v}) + \beta \mathbb{E}_s D(\theta^2 + \lambda, x, s).$$

2. **We show that**

$$D(\theta^2, x, s) \ge \sup_{a \in \tilde{\mathcal{A}}(x,s)} \inf_{\lambda \ge 0} u(x, a, s) + \theta^2(v(x, a, s) + \lambda(v(x, a, s) - \bar{v}) + \beta \mathbb{E}_s D(\theta^2 + \lambda, x, s).$$
(D.4)

By the definition of D, for any $x' \in \mathcal{X}$, $s' \in \mathcal{S}$, $\lambda \ge 0$, there exists $(u'_\lambda(x', s'), v'_\lambda(x', s')) \in \mathcal{U}(x', s')$, s.t. $D(\lambda, x', s') = u'_\lambda(x', s') + \tilde{\lambda} v_\lambda(x', s')$. For any $a \in \tilde{\mathcal{A}}(x, s)$, we have

$$\begin{aligned}
&\inf_{\lambda} u(x, a, s) + \theta^2 v(x, a, s) + \lambda(v(x, a, s) - \bar{v}) + \beta \mathbb{E}_s D(\theta^2 + \lambda, x', s') \\
&= \inf_{\lambda} u(x, a, s) + \theta^2 v(x, a, s) + \lambda(v(x, a, s) - \bar{v}) + \beta \mathbb{E}_s(u'_{\theta_2 + \lambda} + (\theta^2 + \lambda)v'_{\theta_2 + \lambda})) \\
&= \inf_{\lambda} u(x, a, s) + \theta^2 v(x, a, s) + \lambda(v(x, a, s) + \beta \mathbb{E}_s v'_{\theta_2 + \lambda} - \bar{v}) + \beta \mathbb{E}_s(u'_{\theta_2 + \lambda} + \theta^2 v'_{\theta_2 + \lambda})).
\end{aligned}$$

If $(a, u'_{\theta^2}, v'_{\theta^2}) \in \mathcal{F}$, then

$$\begin{aligned}
&\inf_{\lambda} u(x, a, s) + \theta^2 v(x, a, s) + \lambda(v(x, a, s) + \beta \mathbb{E}_s v'_{\theta_2 + \lambda} - \bar{v}) + \beta \mathbb{E}_s(u'_{\theta_2 + \lambda} + \theta^2 v'_{\theta_2 + \lambda})) \\
&\le u(x, a, s) + \theta^2 v(x, a, s) + \beta \mathbb{E}_s(u'_{\theta^2} + \theta^2 v'_{\theta^2}) \le D(\theta^2, x, s).
\end{aligned}$$
(D.5)

If $(a, u'_{\theta^2}, v'_{\theta^2}) \notin \mathcal{F}$ and then one of the following happens:

- For all $s' \in \mathcal{S}$, $(u'(\zeta(x,a,s),s'), v'(\zeta(x,a,s),s')) \in \mathcal{U}(\zeta(x,a,s),s')$, $(a, u', v') \notin \mathcal{F}$. According to the closedness of $\mathcal{U}$, there exists $\underline{L} > 0$, s.t.

$$v(x,a,s) + \beta \mathbb{E}_s v'_{\theta_2 + \lambda} - \bar{v} \leq \underline{L},$$

implying that

$$\inf_\lambda u(x,a,s) + \theta^2 v(x,a,s) + \lambda(v(x,a,s) + \beta \mathbb{E}_s v'_{\theta_2+\lambda} - \bar{v}) + \beta \mathbb{E}_s(u'_{\theta_2+\lambda} + \theta^2 v'_{\theta_2+\lambda})) = -\infty.$$

- We could find a $\lambda^* > 0$, s.t. $v'_{\theta_2 + \lambda^*}$ binds the incentive constraint $v(x,a,s) + \beta \mathbb{E}_s v'_{\theta_2 + \lambda^*} - \bar{v} = 0$ due to the convexity of D. Similar arguments gives the same inequality (D.5).

Therefore,

$$\sup_{a \in \tilde{\mathcal{A}}(x,s)} \inf_\lambda u(x,a,s) + \theta^2 v(x,a,s) + \lambda(v(x,a,s) - \bar{v}) + \beta \mathbb{E}_s D(\theta^2 + \lambda, x', s') \leq D(\theta^2, x, s).$$

We combine **1.** and **2.** and see that the largest fixed point of (D.1) coincides with the largest fixed point of the sup-inf formula.

E   Finite horizon forward looking constraints

In the general formulation, incentive constraints only involve finite periods. In this section, we consider the two-period case. The incentive constraints in the deterministic problems are replaced by

$$\mathbb{E}_{s_t} \sum_{n=0}^1 \beta^n g^i(x_{t+n}, a_{t+n}(s^{t+n}), s_{t+n}) \geq \bar{g}^i, \quad \forall t \in \mathbb{N}, \forall s \in \mathcal{S}^t, \forall i \in \{1, \cdots, I\},$$

and the deterministic problem is[20]

$$\max_{(a_t(s^t)) \in \mathcal{A}^\infty \subset \ell^\infty} \sum_{i=1}^I \gamma^i g^i(x_0, a_0(s^0), s_0) + \mathbb{E}_{s_0} \sum_{t=0}^\infty \beta^t r(x_t, a_t(s^t), s_t)$$

$$\textbf{s.t.} \ \mathbb{E}_{s_t} \sum_{n=0}^1 \beta^n g^i(x_{t+n}, a_{t+n}(s^{t+n}), s_{t+n}) \geq \bar{g}^i, \quad \forall t \in \mathbb{N}, \forall s^t \in \mathcal{S}^t, \forall i \in \{1, \cdots, I\},$$

$$\text{where } x_{t+1} = \zeta(x_t, a_t(s^t), s_t) \text{ and } p(x_t, a_t(s^t), s_t) \geq 0, \quad \forall t \in \mathbb{N}, \ s^t \in \mathcal{S}^t$$

(E.1)

---

[20]Note that the objective function is also different from problem (3.1).

The Lagrangian is then defined as

$$
\begin{aligned}
&L((a_t(s^t)), (\lambda_t^i(h^t)); (\gamma^i), x_0, s_0) \\
&= \sum_{i=1}^{I} \gamma^i g^i(x_0, a_0(s^0), s_0) + \mathbb{E}_{s_0} \sum_{t=0}^{\infty} \beta^t \Bigg(r(x_t, a_t(s^t), s_t) \\
&\quad + \lambda_t^i(s_0, a_0(s_0), \cdots, s_{t-1}, a_{t-1}(s^{t-1}), s_t) \left(\sum_{n=0}^{1} \beta^n g^i(x_{t+n}, a_{t+n}(s^{t+n}), s_{t+n}) - \bar{g}^i \right) \Bigg).
\end{aligned}
\tag{E.2}
$$

Theorem E.1 *For any $x \in \mathcal{X}$, $s \in \mathcal{S}$, $\gamma \in \mathbb{R}_+^I$, the dual value function $D(x, \gamma, s_0)$ defined as*

$$
D(\gamma, x, s_0) := \inf_{(\lambda_t^i(h^t)) \in \Lambda} \sup_{(a_t(s^t)) \in \tilde{\mathcal{A}}^{\infty}} L((a_t(s^t)), (\lambda_t^i(h^t)); (\gamma^i), x_0, s_0)
$$

satisfies the following recursive equation

$$
D(\gamma, x, s) = \inf_{\lambda \in \mathbb{R}_+^I} \sup_{a \in \tilde{\mathcal{A}}(x,s)} \left[\left(r(x,a,s) + \sum_{i=1}^{I} \left(\gamma^i g^i(x,a,s) + \lambda^i(g^i(x,a,s) - \bar{g}^i) \right) \right) + \beta \mathbb{E}_s D(\lambda, x', s') \right],
$$

where $x' = \zeta(x,a,s)$.
$$\tag{E.3}$$

Proof By straightforward algebric simplification, we have

$$
\begin{aligned}
&L((a_t(s^t)), (\lambda_t^i(h^t)); (\gamma^i), x_0, s_0) \\
&= \sum_{i=1}^{I} \gamma^i g^i(x_0, a_0(s^0), s_0) + \mathbb{E}_{s_0} \sum_{t=0}^{\infty} \beta^t \Bigg(r(x_t, a_t(s^t), s_t) \\
&\quad + \sum_{i=1}^{I} \lambda_t^i(s_0, a_0(s_0), \cdots, s_{t-1}, a_{t-1}(s^{t-1}), s_t) \left(\sum_{n=0}^{1} \beta^n g^i(x_{t+n}, a_{t+n}(s^{t+n}), s_{t+n}) - \bar{g}^i \right) \Bigg) \\
&= \left(r(x_0, a_0, s_0) + \sum_{i=1}^{I} \left(\gamma^i g^i(x_0, a_0, s_0) + \lambda_0^i(g^i(x_0, a_0, s_0) - \bar{g}) \right) \right) \\
&\quad + \beta \mathbb{E}_{s_0} \Bigg(\sum_{i=1}^{I} \lambda_0^i g^i(x_1, a_1, s_1) + \mathbb{E}_{s_1} \sum_{t=1}^{\infty} \beta^{t-1} \Bigg(r(x_t, a_t(s^t), s_t) \\
&\quad + \sum_{i=1}^{I} \lambda_t^i(s_0, a_0(s_0), \cdots, s_{t-1}, a_{t-1}(s^{t-1}), s_t) \left(\sum_{n=0}^{1} \beta^n g^i(x_{t+n}, a_{t+n}(s^{t+n}), s_{t+n}) - \bar{g}^i \right) \Bigg) \Bigg).
\end{aligned}
\tag{E.4}
$$

The rest of the proof follows the same procedure as in the proof for Theorem 3.1, and we omit the details here. $\qquad\square$