# Recursive Contracts, Lotteries and Weakly Concave Pareto Sets

Harold Cole
University of Pennsylvania

Felix Kubler
University of Zurich

December 15, 2010

**Abstract**

Marcet and Marimon (1994, revised 1998) developed a recursive saddle point method which can be used to solve dynamic contracting problems that include participation, enforcement and incentive constraints. Their method uses a recursive multiplier to capture implicit prior promises to the agent(s) that were made in order to satisfy earlier instances of these constraints. As a result, their method relies on the invertibility of the derivative of the Pareto frontier and cannot be applied to problems for which this frontier is not strictly concave. In this paper we show how one can extend their method to a weakly concave Pareto frontier by expanding the state space to include the realizations of an end of period lottery over the extreme points of a flat region of the Pareto frontier. With this expansion the basic insight of Marcet and Marimon goes through – one can make the problem recursive in the Lagrangian multiplier which yields significant computational advantages over the conventional approach of using utility as the state variable. The case of a weakly concave Pareto frontier arises naturally in applications where the principal's choice set is not convex but where randomization is possible.

Key Words: Recursive Contracting, Recursive Multipliers, Lotteries.
JEL: C61, C63,

# 1 Introduction

Marcet and Marimon (1994, revised 1998), in their still unpublished research memo, developed a recursive saddle point method which can be used to solve dynamic contracting problems that include participation, enforcement and incentive constraints. Their method uses a recursive multiplier to capture implicit prior promises to the agent(s) that were made in order to satisfy earlier (in time) instances of these constraints. Prior to their work, the standard method to recursively solve dynamic contracting problems treated the promised utility of the agent(s) as a state variable, and used this state variable to capture the implicit prior promises. The main advantage of the recursive multiplier is that it allows the conditional date $t$ problem to be solved without reference to the ex ante date $t$ payoff to the agent(s). In contrast, when one uses ex ante utility as a state variable, there is an overall ex ante condition as to the allocation of utility across states in $t$ that has to hold, and as a result the date $t$ actions and date $t+1$ continuation utilities must be simultaneously solved as a block.

This computational advantage has lead to the recursive multiplier approach being widely used; to cite a few examples Marcet and Marimon (1992), Kahn, King and Wolman (2003), Cooley, Marimon and Quadrini (2004) Attanasio and Rios-Rull (2000), Kehoe and Perri (2002), Aiyagari, Marcet, Sargent and Seppala (2002), Atkeson and Cole (2005), Chien, Cole and Lustig (2009). Despite these applications, there remain fundamental issues with respect to the applicability of these methods. Messner and Pavoni (2004) use a simple example to show that if the Pareto frontier is not strictly concave, then these methods can yield policy functions which are not only suboptimal but infeasible. This is because Marcet and Marimon's method relies on the invertibility of the derivative of the Pareto frontier in order to map the recursive multiplier into the promised utility level. The method must fail in the presence of flat spots on the Pareto frontier. Public randomization, which is commonly used in environments with nonconvex constraint sets, naturally generates these flat spots on the Pareto frontier.[1]

In this paper we show how one can deal with a weakly concave Pareto frontier by expanding the state space to include the realizations of an end of period lottery over the extreme points on the Pareto frontier that share a common slope. The basic idea is as follows. It can be shown that the value of the recursive programming problem in Marcet-Marimon is the same as the optimal value of the contracting problem even if there are flat spots on the frontier. The problem is that if for a given value of the multiplier the frontier is flat, the recursive multiplier approach yields a continuum of current actions which solve the Bellman equation. It is then impossible to pick the correct action that is consistent with previous periods' incentive constraints. Our approach identifies those actions in these sets that yield extreme payoffs. In the case of one agent and one principal these are the action that yields the highest payoff to the agent and the action that yields the highest payoff to the principal, i.e. the two extreme points of the flat spot of the Pareto-frontier. We assume that the principal has access to a public randomization device and can pick a lottery over these two

[1] Lotteries naturally arise in environments with incentive constraints, see for example Prescott and Townsend (1984 A and B), or discrete choices, see for example Rogerson (1988), and Cole and Prescott (1997).

extreme points that satisfies last period's incentive constraints. If in the previous period the multiplier identified a strictly concave region of the frontier, this resulting policy clearly yields the optimal value and is feasible. If in the previous period, the multiplier also pointed to a flat spot on the frontier, one again has to identify the extreme points and continue with this until one is at $t = 0$, or at a strictly concave region of the frontier.

We discuss the fundamental issue that arises with a Pareto frontier that is not strictly concave, and the nature of our solution to this problem in terms of a simple example which we will solve in detail later in the paper. Consider a simple partnership model in which there is a principal and an agent, and the principal must have the participation of the agent in order to run a project which produces a pie of size 1 every period that can be split between the two of them. Each period the principal gets to choose the amount of the pie he eats, $a$, and this implies the amount that the agent eats, $1 - a$. The within period reward functions for the principal is $\log(a)$ and for the agent is $\log(1-a)$. Both are expected utility maximizers and both discount the future at rate $\beta$.

The agent has an initial outside opportunity, and then each period draws an outside opportunity. These opportunities put an initial ex ante lower bound on his payoff, which we call the participation constraint, and a conditional lower bound, which we can an incentive constraint. The ex ante outside opportunity which goes into the participation constraint is $\bar{g}_2 = \frac{1}{1-\beta}\log(1/5)$. Assume that in every period there are two states of the world $s \in \{l, h\}$ which are i.i.d. and equi-probable. The conditional opportunity of the agent depends on the shock, $\bar{g}_1(l) = \frac{1}{1-\beta}\log(\epsilon)$ and $\bar{g}_1(h) = \frac{1}{1-\beta}\log(2/3)$. These opportunities determine the incentive constraint.

We will assume that $a$ is bounded between $\varepsilon$ and $1 - \varepsilon$, where $\varepsilon$ is a small number that serves to bound the payoffs of the agent and the principal. We will consider two cases: (i) the set of actions $a$ is convex and equal to $[\varepsilon, 1-\varepsilon]$, and (ii) the set of actions is not convex. In the nonconvex case, we will allow for public randomization in order to convexify the set of payoffs.

The Pareto frontier, $V(G)$, for this environment are the solutions to the optimal contracting problem for each feasible level of ex ante utility for the agent , $G$, where $V$ is the principal's payoff, and by incentive feasible we mean that the contract satisfies the ex post incentive constraint in every date and state. Because of possibility of public randomization, we let $\psi(a, s)$ denote a probability distribution over the possible actions, and implicitly define the Pareto frontier in the following functional equation, which we refer to as the Pareto problem,

$$V(G) = \max_{\psi(a,s),G(s)} E\left\{\int_a \log(a)d\psi(a,s) + \beta V(G(s))\right\}$$

subject to the ex post incentive constraint (IC)

$$\log(1-a) + \beta G(s) \geq \bar{g}_1(s) \text{ for all } a \text{ in the support of } \psi \text{ and each } s,$$

and the ex ante participation constraint (PC)

$$E\left\{\int_a \log(1-a)d\psi(a,s) + \beta G(s)\right\} \geq G.$$

The solution to the original contracting problem is generated when we set $G = \bar{g}_2$. If the set of possible actions $a$ by the principal is convex, then randomization is degenerate and one can show that the Pareto frontier is strictly concave. However if the action set is not convex, then public randomization may be necessary and as a result the Pareto frontier is weakly concave but not strictly so. It is now easy to explain that in this case the multiplier approach in Marcet and Marimon need not yield the optimal policy.

We will make the following assumptions about the solution to this problem at $G = \bar{g}_2$ (i.e. when we impose the participation constraint). Assume that the IC binds in state $s = h$. Assume that the PC binds and that its shadow value is $\gamma$. Clearly the IC cannot bind in state $s = l$ since the bound is the lowest feasible transfer to the agent. If we denote the multiplier on the IC constraint by $\lambda(s)$.

If the action set is convex there will be a unique action in the support of $\psi$, $a(s)$, and $V(G)$ is strictly concave. Moreover, at the solution

$$\frac{1}{a(l)} = \gamma \frac{1}{1 - a(l)},$$

$$V'(G(l)) = \gamma,$$

$$\frac{1}{a(h)} = (\gamma + \lambda(h)) \frac{1}{1 - a(h)},$$

and

$$V'(G(h)) = -(\gamma + \lambda(h)).$$

And, these conditions uniquely determined the solution given the multipliers $\gamma$, $\lambda(l)$ and $\lambda(h)$.

Now consider the *Social Planning problem* associated with this Pareto problem. Recall that the Pareto-frontier is described by $V(G)$ and that $\gamma$ is the welfare weight for the principal that is implied by the participation constraint. The social planning problem is then given by

$$\max_{\psi(a,s),G(s)} E_s \left\{ \int_a \log(a) d\psi(a,s) + \beta V(G(s)) + \gamma \left[ \int_a \log(1-a) d\psi(a,s) + \beta G(s) \right] \right\}$$

subject to the incentive constraint. In the case in which the action set is convex, and hence the Pareto frontier is strictly concave, the solution to this problem is unique and will satisfy the same optimality conditions as the Pareto problem with $G = \bar{g}_2$. Hence, the solution to the social planning problem will correspond to the solution to the planning problem, and the payoff to the agent is uniquely determined by his weight in the social planning problem.

In the case in which the action set is not convex, and hence the Pareto frontier is not necessarily strictly concave, this correspondence need not hold precisely because the payoff to the agent is not uniquely determined by his social planning weight. For example, assume that in the solution to the original problem $G(l)$ lay on a linear portion of the Pareto frontier. It follows then that the value of the social planning problem doesn't change if we vary $G(l)$ along its linear portion since the $V$ and $G$ trade-off is exactly $-\gamma$, which is the weight in the social planning problem. Or, for another example, assume that in the original problem the randomization in the low state was not degenerate,and we were randomizing over $a_1$ and $a_2$. Then, it must be the case that

$$\log(a_1) + \gamma \log(1 - a_1) = \log(a_2) + \gamma \log(1 - a_2),$$

and any randomization over these two points can be part of a solution to the social planning problem. Hence, the payoff to the agent is not uniquely determined by his weight in the social planning problem. These example illustrate than when the Pareto frontier is not strictly concave, there can be solutions to the social planning problem in which the PC constraint is violated or is slack. Of course the solution set will also include the solution to the planning problem.

If we consider the Lagrangian associated with our Social Planning problem,

$$\max_{\psi(a,s),G(s)} E \left\{ \begin{array}{c} \int_a \log(a) d\psi(a,s) + \beta V(G(s)) + \gamma \left[\int_a \log(1-a) d\psi(a,s) + \beta G(s)\right] \\ \int_a \lambda(a,s) \left[\log(1-a) + \beta G(s) - \bar{g}_1(s)\right] d\psi(a,s) \end{array} \right\}$$

these same issues will arise. The Lagrangian for the social planning problem will pick out the correct $[\psi(a,h), G(h)]$ because of the requirement that the IC hold with equality if $\lambda(h) > 0$, which it is. However the PC may not hold with equality in the solution to the either the Social Planning problem or its associated Lagrangian problem.

A recursive saddlepoint methodology uses the sum of the past multipliers on the PC and IC constraints in a dynamic optimization problem as the state variable to determine the relative treatment of the principal and the agent. Now the problem becomes even worse. In the solution to the social planning problem $V'(G(s)) = -\left[\gamma + \lambda(s)\right]$. Since both $G(l)$ and $G(h)$ could fall on on linear portions of the Pareto frontier, this condition is not sufficient to uniquely determine either of these continuation payoffs for the agent. Hence it cannot ensure that either the PC or the IC conditions hold as equalities which they must in the true solution given our assumptions. Of course the set of possible solutions to this problem will contain the true solution. However, since the value $V(G) + [\gamma + \lambda(s)]\, G$ is the same for all $G$ such that $V'(G(s)) = -\left[\gamma + \lambda(s)\right]$, the lack of strict concavity of the Pareto Problem is a problem for the policy functions and not the value of the solution.

One simple way around this problem is to appropriately randomize over extreme points in the solution set to the social planning problem. To illustrate how this works, let $A(s)$ denote the set of $\{a(s), G(s)\}$ pairs that are solutions to

$$\max_{a(s),G(s)} \log(a(s)) + \beta V(G(s)) + (\gamma + \lambda(s)) \left[\log(1-a) + \beta G(s)\right].$$

Let

$$\begin{aligned} G^H(s) &= \max_{\{a,G\} \in A(s)} \log(1-a) + \beta G \\ G^L(s) &= \min_{\{a,G\} \in A(s)} \log(1-a) + \beta G, \end{aligned}$$

Let $\pi(s)$ denote the conditional probability of the agent getting the lowest optimal payoff and $1 - \pi(s)$ of getting the highest payoff. In this case conditional expected the payoff of the agent is

$$\pi(s) G^L(s) + (1 - \pi(s)) G^H(s),$$

and the principal is

$$\pi(s)V(G^L(s)) + (1-\pi(s))V(G^H(s)).$$

By choosing $\pi(l)$ and $\pi(h)$ appropriately, we can ensure that these conditional payoffs are such that both the PC and the IC hold as equalities. Under this resolution the state space tomorrow becomes $(s', \gamma + \lambda(s), \pi(s))$, where $s'$ is tomorrow state, and this is sufficient to generate a correct policy choice.

In the dynamic problem that we will consider below this randomization is with respect to the selection of the continuation payoff and hence can be done at the end of the period. As a result the state space that we will use will be even simpler since it will be $(s', \gamma + \lambda, i = L$ or $H)$; that is, we will include an indicator of which element of the extreme points in the solution set will be used.

The rest of the paper is organized as follows. In Section 2 we describe the general contracting problem. In Section 3 we explain the recursive multiplier approach from Marcet and Marimon and give a simple example to illustrate that without strict concavity this approach cannot yield the correct policy. In Section 4 we describe our solution to the problem and link it to the promised utility approach. Section 5 revisits the example from the introduction. In Section 6 we argue that more complicated problems can easily be solved numerically with our method.

# 2 The general problem

Assume the exogenous shock $(s_t)$ follows a finite Markov chain with transition $\Pi$ and support $\mathcal{S} = \{1, ..., S\}$. Each period the principal randomizes over a compact action set $\mathcal{A} \subset \mathbb{R}^n$ – to simplify the notation and in anticipation of our results below we assume without loss of generality that he chooses a *simple* probability distribution (i.e. a distribution with finite support) $\psi_t : \mathcal{A} \to \mathbb{R}_+$, $\psi_t \in \Psi$. The choice can depend on the history of realized shocks and realized actions, i.e. $\psi_t = \psi(h^{t-1}, s_t)$ where $h^{t-1} = (s_0, a_0, \ldots s_{t-1}, a_{t-1})$. In a slight abuse of notation we write $\Psi^\infty$ to denote the space of all (history-dependent) sequences. We denote the support of $\psi_t$ by $\text{supp}(\psi_t) = \{a \in \mathcal{A} : \psi_t(a) > 0\}$. Note that the space of all simple probability distribution forms a vector space and convexity is well defined.

It is useful to denote by $\Pi_t^\psi(h^{t+n})$ the conditional probability of history $h^{t+n}$, given history $h^t$ and a choice $(\psi_t) \in \Psi^\infty$. We assume that there is a fixed initial probability distribution over the shock in period 0, $s_0$ and $E_{-1}$ denotes the expectation under this distribution. We write the expectation of a function $f(.)$ that depends on the realized $a_{t+n}$ as

$$E_t^\psi(f(a_{t+n})) = \sum_{h^{t+n} \in \text{supp}(\Pi_t^\psi)} \Pi_t^\Psi(h^{t+n}) f(a_{t+n}(h)).$$

We use $E_{t-1,s_t}$ for the expectation conditional on a history $h^{t-1}$ and a realized shock $s_t$ – $E_{s_0}$ denotes the expectation conditional on a realized first period shock, $s_0$.

The physical state, $x$, realizes in a compact subset of Euclidean space $\mathcal{X} \subset \mathbb{R}^m$ and depends on last periods' realized action, last periods's state and last period's realized shock but not on the current shock (for our approach, it is trivial to also let the state depend on the current shock, but this adds little to the economics of the model), i.e. the law of motion for $x_t$ is given by $x_{t+1} = \zeta(x_t, a_t, s_t)$.

There are $I$ agents, $i = 1, ..., I$. The principal can randomize over the actions, but for each agent $i = 1, ..., I$ the incentive constraint ($IC_i$) is assumed to hold conditional on each realized action. The participation constraint ($PC_i$), on the other hand, only has to hold ex ante, before either shock or actions are realized.

Our contracting problem then can be written in sequence form as

$$\max_{(\psi_t)\in\Psi^\infty} E_{-1}^{\psi} \sum_{t=0}^{\infty} \beta^t r(x_t, a_t, s_t) \text{ subject to} \tag{1}$$

$$x_{t+1} = \zeta(x_t, a_t, s_t) \tag{2}$$

$$p(x_t, a_t, s_t) \geq 0, \quad \forall a_t \in \text{supp}(\psi_t) \tag{3}$$

$$g^i(x_t, a_t, s_t) + E_t^{\psi} \sum_{n=1}^{\infty} \beta^n g^i(x_{t+n}, a_{t+n}, s_{t+n}) \geq \bar{g}_1^i, \ \forall i = 1, ..., I, \ \forall a_t \in \text{supp}(\psi_t) \tag{4}$$

$$E_{-1}^{\psi} \sum_{t=0}^{\infty} \beta^t g^i(x_t, a_t, s_t) \geq \bar{g}_2^i, \forall i = 1, ..., I \tag{5}$$

$$x_0 \text{ given .} \tag{6}$$

For consistency, we assume that for all $(x, a, s) \in \mathcal{X} \times \mathcal{A} \times \mathcal{S}$, $\zeta(x, a, s) \in \mathcal{X}$ whenever $p(x, a, s) \geq 0$. It is useful to combine constraints (2) and (3) – for a fixed $x_0$, we write $x_t = x(h^{t-1})$ to denote the state at $t$ resulting from history $h^{t-1}$ as well as $\mathcal{A}(h^{t-1}, s_t)$ to be the set of all actions satisfying (3) given $x(h^{t-1})$ and the current shock $s_t$. Since the space of exogenous shocks is assumed to be finite, we frequently write $r_s(x, a)$ instead of $r(x, a, s)$ and $g_s^i(x, a)$ instead of $g^i(x, a, s)$. To simplify notation we sometimes use $g^0$ to denote the reward function of the principal, i.e. $g^0(x, a, s) \equiv r(x, a, s)$. In general we could let the outside option in the (IC) constraint, $\bar{g}_1^i$ depend on the shock $s$. It will be easy to see that this does not change our analysis, we just drop the dependence to simplify notation.

We make the following assumptions on the fundamentals.

**Assumption 1**

1. *The reward functions $r_s : \mathcal{X} \times \mathcal{A} \to \mathbb{R}$ and the constraint functions $g_s^i : \mathcal{X} \times \mathcal{A} \to \mathbb{R}$ are continuous and bounded (both from below and above) for all shocks $s \in \mathcal{S}$.*

2. *There is discounting, i.e. $\beta \in (0, 1)$.*

3. *There is an $\epsilon > 0$ such that for each initial condition $(x_0) \in \mathcal{X}$ there exists an action $(\psi_t)$ such that for all agents $i = 1, ..., I$,*

$$\inf_{h^{t-1}, s_t, a_t \in supp(\psi_t)} \left[ g^i(x_t, a, s_t) + E_t^{\psi} \sum_{n=1}^{\infty} \beta^n g^i(x_{t+n}, a_{t+n}, s_{t+n}) - \bar{g}_1^i \right] > \epsilon$$

*and*

$$E_{-1}^{\psi} \sum_{t=0}^{\infty} \beta^t g^i(x_t, a_t, s_t) - \bar{g}_2^i > \epsilon.$$

Assumption 1.1 is perhaps stronger than needed. Assumption 1.2 and the Slater condition 1.3 are standard. Condition 1.3 can of course be rewritten in terms of non-randomized actions.

# 3 The social planning problem

In this section, we disregard the participation constraint (5) and focus on a social planner's problem for given welfare weights. Without the participation constraint the planner's problem can be divided into $S$ subproblems, one for each initial shock. To solve the original problem the welfare weight $\gamma$ needs to be chosen to ensure that this constraint holds, i.e. becomes the initial multiplier on the participation constraint.

## 3.1 Non-recursive formulation

For a given choice of $(\gamma_1, \ldots, \gamma_I) \in \Gamma = [0, \infty)^I$, for a given $s_0$ and for an admissible state $x_0 \in \mathcal{X}$, define the social planner's problem (in sequence form) as

$$\max_{(\psi_t) \in \Psi^\infty} E^\psi_{s_0} \left[ \sum_{t=0}^{\infty} \beta^t \left( r(x_t, a_t, s_t) + \sum_{i=1}^{I} \gamma_i g^i(x_t, a_t, s_t) \right) \right] \text{ subject to}$$

$$\text{supp}(\psi_t) \subset \mathcal{A}(h^{t-1}, s_t) \text{ for all } h^{t-1}, s_t \tag{7}$$

$$g^i(x_t, a_t, s_t) + E^\psi_t \sum_{n=1}^{\infty} \beta^n g^i(x_{t+n}, a_{t+n}, s_{t+n}) \geq \bar{g}^i_1 \quad \forall a_t \in \text{supp}(\psi_t), \ \forall t, \ \forall i$$

### 3.1.1 The Lagrangian

It is useful to introduce the Lagrangian for this problem and to write

$$L\left((\lambda^i_t), (\psi_t); (\gamma_i), x_0, s_0\right) = E^\psi_0 \sum_{t=0}^{\infty} \beta^t \tag{8}$$

$$\left( r(x_t, a_t, s_t) + \sum_{i=1}^{I} \left( \gamma_i g^i(x_t, a_t, s_t) + \lambda^i_{t,a_t} \left(\sum_{n=0}^{\infty} \beta^n g(x_{t+n}, a_{t+n}, s_{t+n}) - \bar{g}^i_1\right) \right) \right).$$

Note that for given $(\lambda_t)$ this function is concave in $\psi$ and the constraint set $\{(\psi_t) : \text{supp}(\psi_t) \subset \mathcal{A}(h^{t-1}, s_t) \text{ for all } h^{t-1}, s_t\}$ is a convex set. Therefore, it is standard to show (e.g. Luenberger (1969, Theorem 2, page 221)) that if there exist $(\psi_t)^*$ with $\text{supp}(\psi^*_t) \subset \mathcal{A}(h^{t-1}, s_t)$ for all $h^{t-1}, s_t$ and a $(\lambda_t)^* \geq 0$ that satisfy

$$L\left((\lambda_t), (\psi_t)^*\right) \geq L\left((\lambda_t)^*, (\psi_t)^*\right) \geq L\left((\lambda_t)^*, (\psi_t)\right) \tag{9}$$

for all $(\lambda_t) \geq 0$ and all $(\psi_t)$ with $\text{supp}(\psi_t) \subset \mathcal{A}(h^{t-1}, s_t)$ for all $h^{t-1}, s_t$ then $(\psi_t)^*$ is a solution to the original problem (7)). The results in Rustichini (1998) and Dechert (1982) imply that under our assumptions (1)-(3) there exists such a saddle point with the sequences $(\lambda_t)^*$ being bounded.

## 3.2 Recursive multiplier formulation

We now define a functional equation similar to the one in Marcet and Marimon (1994) and prove that the value-function of the problem gives the value of the social planner's problem. The following functional equation defines the social planning problem in recursive multiplier form

$$
\begin{aligned}
&F(\gamma, x, s) = \sup_{\psi \in \Psi} \inf_{\lambda \geq 0} \sum_{a \in \text{supp}(\psi)} \psi(a) \\
&\left[ r(x, a, s) + \sum_{i=1}^{I} \left( \gamma^i g^i(x, a, s) + \lambda_a^i (g^i(x, a, s) - \bar{g}_1^i) \right) + \beta E_s F(\gamma + \lambda_a, x', s') \right] \\
&\quad \text{subject to} \\
&x' = \zeta(x, a, s) \\
&p(x, a, s) \geq 0, \quad \forall a \in \text{supp}(\psi).
\end{aligned} \tag{10}
$$

The following theorem guarantees the existence of a unique solution.

**Theorem 1** *The functional equation (10) has a unique solution in the space of bounded functions $F_s : [0, \infty) \times \mathcal{X} \to \mathbb{R}$, $s \in \mathcal{S}$.*

The main problem in proving the theorem is to bound the multipliers $\lambda$. For this we need to first consider a slightly different problem: For a given $\bar{\lambda}$ and for bounded functions $f : ([0, \infty) \times \mathcal{X})^S \to \mathbb{R}$ define the operator

$$
\begin{aligned}
&T_{\bar{\lambda}}(f) = \max_{\psi \in \Psi} \min_{\lambda \in [0, \bar{\lambda}]^I} \sum_{a \in \text{supp}(\psi)} \psi(a) \\
&\left[ r(x, a, s) + \sum_{i} \left( \gamma^i g^i(x, a, s) + \lambda_a^i (g^i(x, a, s) - \bar{g}_1^i) \right) + \beta E_s f(\gamma + \lambda_a, x') \right] \\
&\quad \text{subject to} \\
&x' = \zeta(x, a, s) \\
&p(x, a, s) \geq 0, \quad \forall a \in \text{supp}(\psi),
\end{aligned}
$$

where $E_s f = \sum_{s'} \pi(s, s') f_{s'}$. By Assumption 1.1. and for a given $\bar{\lambda}$, clearly $T_{\bar{\lambda}}$ maps bounded functions into bounded functions.

Therefore it is easy to proof the following lemma.

**Lemma 2** *The operator $T_{\bar{\lambda}}$ has a unique fixed point in the (Banach) space of bounded functions.*

**Proof.** We apply the contraction mapping theorem and verify Blackwell's sufficient conditions monotonicity (M) and discounting (D) (see e.g. Stokey and Lucas (1989), Theorem 3.3).

(M) Given bounded $f^1$ and $f^2$ with $f^1(\gamma, x) \leq f^2(\gamma, x)$ for all $(\gamma, x)$, let $(\psi^1, \lambda^1)$ be the max-minimizer of $T_{\bar{\lambda}} f^1(\gamma, x)$ Clearly,

$$T_{\bar{\lambda}} f^1(\gamma, x) \leq \min_{\lambda \in [0,\bar{\lambda}]^I} \sum_{a \in \operatorname{supp}(\psi^1)} \psi^1(a)$$
$$\left[ r(x,a,s) + \sum_i \left( \gamma^i g^i(x,a,s) + \lambda_a^i (g^i(x,a,s) - \bar{g}_1^i) \right) + \beta E_s f^2(\gamma + \lambda_a, x') \right]$$
$$\leq T_{\bar{\lambda}} f^2(\gamma, x).$$

(D) Just substituting in yields $T_{\bar{\lambda}}(f + a) = T_{\bar{\lambda}}(f(\gamma, x) + a) = T_{\bar{\lambda}} f + \beta a$. $\Box$

We can now prove the theorem by letting $\bar{\lambda}$ become sufficiently large and showing that the constraint never binds.

**Proof of Theorem 1** Since the problem in the functional Equation (10) is convex, it suffices to show that for sufficiently large $\bar{\lambda}$ the constraint $\lambda_a(h^t) \leq \bar{\lambda}$ can never be binding. Let $\epsilon$ be as in Assumption 1.3. Define $d = \frac{1}{1-\beta} \left( \sup_{x,a,s} r(x,a,s) - \inf_{x,a,s} r(x,a,s) \right)$ and let $\bar{\lambda} = 2\frac{d}{\epsilon}$. Suppose at some state $(x, s) \in \mathcal{X} \times \mathcal{S}$ and some $\gamma \geq 0$ the constraint is binding. By Assumption 1.3 there exists a policy which yields a total reward of at least $2d$. But for $\lambda = 0$, the value of total reward is smaller and hence $\bar{\lambda}$ cannot be the minimizing choice of $\lambda$. $\Box$

Having established the existence of a solution of the functional equation (10) raises the question how this solution relates to the solution of the original problem. This turns out to be somewhat intricate.

## 3.3 Policy- and value-functions

Unfortunately, while the value function $F$ is well defined and continuous, the set of arguments that maximize $F$ might contain infinitely many solutions. While it is true that if a solution to original saddle point problem exists that this also solves (9), the converse is false, in the sense that there can be many solutions to (9) that do not solve the saddle point problem. This was first observed by Messner and Pavoni (2004). The problem has nothing to do with infinite horizon or uncertainty but is simply caused by the fact that one cannot recover policies from correspondences in a min-max problem. One can illustrate the problem with a trivial two period example.

### 3.3.1 Simple example

Suppose we want to solve the following problem

$$V = \max_{(a_1,a_2) \in [0,1]^2} -a_1 - a_2 \text{ subject to}$$
$$a_1 + a_2 \geq 1$$

The value of the corresponding Lagrangian, can be obtained recursively (or by backward induction). We can write

$$\begin{aligned} V_2(\lambda) &= \max_{a_2 \in [0,1]} (-a_2 + \lambda a_2) \\ V_1 &= \max_{a_1 \in [0,1]} \min_{\lambda \geq 0} (-a_1 + \lambda(a_1 - 1) + V_2(\lambda)) \end{aligned}$$

and get the correct value of the saddle point

$$V_1 = \max_{(a_1,a_2) \in [0,1]^2} \min_{\lambda \geq 0} (-a_1 - a_2) + \lambda(a_1 + a_2 - 1).$$

However, one cannot recover the correct policies from this if one takes the Lagrange multiplier as a state variable: The correspondence solving the problem in the second period is given by

$$a_2(\lambda) = \arg \max_{a_2 \in [0,1]} (-a_2 + \lambda a_2) = \begin{cases} 0 & \lambda < 1 \\ [0,1] & \lambda = 1 \\ 1 & \lambda > 1. \end{cases}$$

So clearly, one element of the recursive problem's argmax is $a_1 = 1/2$, $\lambda = 1$ and $a_2 = 0$. But this is not a feasible point. Another element of the recursive problem's solution is $a_1 = 1/2, \lambda = 1$ and $a_2 = 1$. This is a feasible point, but obviously suboptimal.

So clearly, the recursive formulation has solutions which are infeasible and has solutions that are feasible but suboptimal. However note that in the above problem the solution to the optimization problem is in the correspondence of possible solutions to the recursive formulations.

A somewhat unrelated notational issue concerns how to properly define the arguments of a max-min problem. If we consider the original problem

$$\max_{(a_1,a_2) \in [0,1]^2} \min_{\lambda \geq 0} (-a_1 - a_2) + \lambda(a_1 + a_2 - 1),$$

then in the solution $\lambda$ is implicitly a function of the action $(a_1, a_2)$, where this function is

$$\lambda(a) = \begin{cases} \infty \text{ if } a_1 + a_2 < 1 \\ 1 \text{ if } a_1 + a_2 = 1 \\ 0 \text{ if } a_1 + a_2 > 1 \end{cases},$$

and this function serves to enforce the constraint. In this case, $a_1 \in [0, 1]$ and $a_2 = 1 - a_1$ is a solution to this problem and the value of the multiplier at the solution is of course $\lambda = 1$. However, if the we consider the min-max version of this problem,

$$\min_{\lambda \geq 0} \max_{(a_1,a_2) \in [0,1]^2} (-a_1 - a_2) + \lambda(a_1 + a_2 - 1),$$

then $\lambda = 1$, $a_1 \in [0, 1]$ and $a_2 = 1 - a_1$ is a solution to this problem but it is not the only one. In particular, with $\lambda = 1$ any choices of $a_1$ and $a_2$ are solutions to the problem, including $a_1 = a_2 = 1$ and $a_1 = a_2 = 0$. Hence, one cannot enforce the constraint in a min-max version of the problem. With convexity, we can apply the max-min theorem and the value of the two problems is identical, however the solution to the first problem is the 'correct' solution. It will therefore be useful to make the following definition. We say that $(\bar{x}, \bar{y}) \in \arg\max_x \min_y f(x,y)$ if $\bar{x} \in \arg\max_x[\min_y f(x,y)]$, if $\bar{y} \in \arg\min_y f(\bar{x}, y)$ and if $\bar{y} \in \arg\min_y[\max_x f(x,y)]$.

Both of these issues arise because of the lack of strict concavity our problem. If the solution is unique, then neither is a problem.

## 3.4 Principle of optimality

We are given a solution $F^*$ to the functional equation (10). In order to construct a policy which solves the saddle point problem (and therefore the original maximization problem), using the above convention about the arg max min, we can define the correspondence

$$
\begin{aligned}
& C(\gamma, x, s) = \arg\max_{\psi} \min_{\lambda} \sum_{a \in \mathrm{supp}(\psi)} \psi(a) \\
& \left[ r(x,a,s) + \sum_{i=1}^{I} \left( \gamma^i g^i(x,a,s) + \lambda_a (g^i(x,a,s) - \bar{g}_1^i) \right) + \beta E_s F^*(\gamma + \lambda_a, x', s') \right] \\
& \quad \text{subject to} \\
& x' = \zeta(x,a,s) \\
& p(x,a,s) \geq 0, \quad \forall a \in \mathrm{supp}(\psi).
\end{aligned}
$$

In this subsection, we want to prove the following theorem.

**Theorem 3** *Given solutions to the recursive problem, $(F^*, C)$, there exists a sequence $(\psi^*(h^t), \lambda^*(h^t))$ with $(\psi^*(h^t), \lambda^*(h^t)) \in C(\gamma(\lambda_0, ..., \lambda_{t-1}), x(h^t), s_t)$ for all $h^t$ that solves the saddle point problem (9). Moreover, the value of the saddle point problem at $(\gamma, x, s)$ is given by $F^*(\gamma, x, s)$.*

The proof quite similar to Bellman's principle of optimality (see e.g. Stokey and Lucas (1989), Section 4.1).

**Proof.** We first proof that if there exists a solution to the saddle point problem, the value must coincide with $F^*$. For this define

$$
\begin{aligned}
& F_\tau(\gamma_0, x_0, s_0) = \sup_{\psi_0, \ldots, \psi_\tau} \inf_{\lambda_0, \ldots, \lambda_\tau} E_0^{\psi_0, \ldots, \psi_\tau} \sum_{t=0}^{\tau} \beta^t \\
& \left( r(x_t, a_t, s_t) + \sum_i \left( \lambda_{t,a_t}^i (g^i(x_t, a_t, s_t) - \bar{g}_1^i) + g^i(x_t, a_t, s_t)(\gamma_0^i + \sum_{n=0}^{t-1} \lambda_n^i) \right) \right) \\
& + E_0^{\psi_0, \ldots, \psi_\tau} \beta^{\tau+1} F^* \left( \left( \gamma_0 + \sum_{n=0}^{\tau} \lambda_n \right), x_{\tau+1}, s_{\tau+1} \right) \\
& \quad \text{subject to } \mathrm{supp}(\psi_t) \subset \mathcal{A}(h^t) \text{ for all } h^t, \quad x_0, s_0 \text{ given.}
\end{aligned}
$$

Note that for $\tau = 1$ we obtain

$$F_1(\gamma_0, x_0, s_0) = \sup_{\psi_0,\psi_1} \inf_{\lambda_0,\lambda_1} E_0 \sum_{t=0}^{1} \beta^t \sum_{a\in \text{supp}(\psi_t)} \psi_t(a)$$

$$\left( r(x_t, a, s_t) + \sum_i \left( \lambda^i_{a,t}(g^i(x_t, a, s_t) - \bar{g}^i) + g^i(x_t, a, s_t)(\gamma^i_0 + \sum_{n=0}^{t-1} \lambda^i_n) \right) \right)$$

$$+E_0\beta^2 F^* \left( \left( \gamma_0 + \sum_{n=0}^{1} \lambda_n \right), x(h^3), s_3 \right)$$

subject to $\text{supp}(\psi_t) \subset \mathcal{A}(h^t)$ for all $h^t$, $\quad x_0, s_0$ given.

By the Fan's minmax theorem and since the problem is additively separable this can be rewritten as

$$F_1(\gamma_0, x_0, s_0) = \sup_{\psi_0} \inf_{\lambda_0} E_0 \sum_{a\in \text{supp}(\psi_0)} \psi_0(a)$$

$$\left( r(x_0, a, s_0) + \sum_i \left( \lambda^i_{0a}(g^i(x_0, a, s_0) - \bar{g}^i_1 \right) + g^i(x_0, a, s_t)\gamma^i_0) \right)$$

$$+\beta E_0 \sup_{\psi_1} \inf_{\lambda_1} \sum_{a\in \text{supp}(\psi_1)} \psi_1(a)$$

$$\left( r(x_1, a, s_1) + \sum_i \left( \lambda^i_{1a} \left( g^i(x_1, a, s_1) - \bar{g}^i_1 \right) + g^i(x_1, a, s_1)(\gamma^i_0 + \lambda^i_0) \right) \right)$$

$$+E_0\beta^2 F^* \left( \left( \gamma_0 + \sum_{n=0}^{1} \lambda_n \right), x(h^2), s_3 \right)$$

subject to $\text{supp}(\psi_t) \subset \mathcal{A}(h^t)$ for all $h^t$, $\quad x_0, s_0$ given.

Clearly $F_1 = F^*$ and in fact by induction, $F_\tau(\gamma_0, x_0, S_0) = F^*(\gamma_0, s_0, s_0)$ for each $\tau = 1, 2, ...$
.

On the other hand, we can write

$$F_\tau(\gamma_0, x_0, s_0) = \sup_{\psi_0,\ldots,\psi_\tau} \inf_{\lambda_0,\ldots,\lambda_\tau} E_0^{\psi_0,\ldots,\psi_\tau} \sum_{t=0}^{\tau} \beta^t$$

$$\left( r(x_t, a_t, s_t) + \sum_{i=1}^{I} \left( \gamma^i g^i(x_t, a_t, s_t) + \lambda^i_{t,a_t}(\sum_{n=0}^{\tau-t} \beta^n g^i(x_{t+n}, a_{t+n}, s_{t+n}) - \bar{g}^i_1) \right) \right) +$$

$$E_0^{\psi_0,\ldots,\psi_\tau} \beta^{\tau+1} F^* \left( \left( \gamma_0 + \sum_{n=0}^{\tau} \lambda_n \right), x_{\tau+1}, s_{\tau+1} \right).$$

Recall that the value of the saddle point is

$$L^*(\gamma_0, x_0, s_0) = \sup_{(\psi_t)} \inf_{(\lambda_t)} E_0^{\psi} \sum_{t=0}^{\infty} \beta^t \tag{11}$$

$$\left( r(x_t, a_t, s_t) + \sum_{i=1}^{I} \left( \gamma_i g^i(x_t, a_t, s_t) + \lambda^i_{t,a_t} (\sum_{n=0}^{\infty} \beta^n g(x_{t+n}, a_{t+n}, s_{t+n}) - \bar{g}^i_1) \right) \right).$$

Since $F^*$ is bounded, the reward functions are bounded and since the optimal multipliers to the saddle point problem are bounded (see Rustichini (1998)) it follows that

$$\lim_{\tau \to \infty} L^*(\gamma_0, x_0, s_0) - F_\tau(\gamma_0, x_0, s_0) = 0.$$

Therefore the value of the saddle point must be equal to $F^*$.

Furthermore the solution to the saddle point equation is certainly feasible for the recursive social planner problem. Since it gives the optimal value, $F^*$, it must satisfy $(\psi_t^*, \lambda_t^*) \in C(\gamma(\lambda_0^*, ..., \lambda_{t-1}^*), x(h^t), s_t)$ for all $t$. $\square$

# 4 Recovering policy-functions

The previous theorem shows that among the (possibly many) solutions to the functional equation (10) there is one which also solves our original problem. The question still remains how to find it. In this section, we explicitly construct a solution to the maximization problem, using the solution to the functional equation (10).

We proceed in two steps. First we consider the general problem, here it will become clear that we are mixing the promised utility approach with the recursive multiplier approach of Marcet and Marimon (1994) to obtain well defined policy functions[2]. We then make an assumption on the solution set of the problem which simplifies the analysis considerably and allows us to derive a simple characterization of the solution.

## 4.1 The general approach

By the maximum principle it is easy to see that the correspondence $C(.)$, as defined in the previous section, is upper hemi-continuous. We can therefore define

$$A(\gamma, x, s) = \{(a, \lambda) : \ \exists (\psi, \lambda) \in C(\gamma, x, s) : a \in \text{supp}(\psi)\}.$$

Note that all $(a, \lambda) \in A(\gamma, x, s)$ yield the same value of the objective function (of the social planner's problem, given $\gamma$). Note also that if $(\psi_t, \lambda_t)_{t=0}^{\infty}$ solves the saddle point problem, then in fact each $a_0 \in \text{supp}(\psi_0)$ is feasible and optimal. As the example in Section 3.3.1 shows, the problem comes about because one needs to establish a link between periods. It is well known that one way to do this is to carry agents' promised utilities.

[2] In Section 4 below we make explicit how our approach compares to the promised utility approach.

For a given 'state[3]' $(\gamma, x, s)$, there might be several possible utilities that can be generated by the several possible actions $(a, \lambda) \in A(\gamma, x, s)$ and by the randomizations over these actions. In a slight abuse of notation, we use $G$ both for the function and the correspondence of possible utility levels and collect these in a set $G(\gamma, x, s)$. The correspondence, $G$, from the state to the set of possible (i.e. feasible) utilities, is defined by the following system.

$$
\begin{aligned}
G(\gamma, x, s) = \text{conv}\{ \quad & (G^0, G^1, \dots, G^I) \in \mathbb{R}^{I+1} : \exists (a, \lambda) \in A(\gamma, x, s) \\
& \exists (G_{s'}^{0\prime}, \dots, G_{s'}^{I\prime}) \in G(\gamma + \lambda, \zeta(x, a, s), s') \text{ for all } s' \\
& G^0 = r(x, a, s) + \beta E_s G^{0\prime} \\
& G^1 = g^1(x, a, s) + \beta E_s G^{1\prime}, G^1 \geq \bar{g}_1^1 \\
& \vdots \\
& G^I = g^I(x, a, s) + \beta E_s G^{I\prime}, G^I \geq \bar{g}_1^I \},
\end{aligned}
$$

were conv$(A)$ denotes the convex hull of a set $A$.

For each admissible state, $\gamma, x, s$, the set $G(\gamma, x, s)$ is a non-empty and convex subset of the $I$-dimensional hyperplane with normal-vector $(1, \gamma_1, \dots, \gamma_I)$. That is $G^0 + \sum_{i=1}^I \gamma_i G^I = const.$ for all $(G^0, \dots, G^I) \in G(\gamma, x, s)$. If the problem is strictly concave, this set will contain only one point – in general, however, it has 'full' dimension $I$.

Obviously there must be a (possible set valued map) from $(\gamma, x, s)$ and a given $G \in G(\gamma, x, s)$ to randomized actions, $\psi$, multipliers $\lambda$, and next period's utilities $(G'_{s'})_{s'=1\dots S}$. Note that by Caratheodory's theorem (Caratheodory (1911)), each point in the set can be obtained by randomizing over at most $I+1$ actions $a$. If we define the state-space to contain utilities promised in the last period as well as $\gamma, x$ and $s$, we can easily construct policies that are optimal and feasible. Given an enlarged state $(\gamma, x, s)$ and $G \in G(\gamma, x, s)$ we now say that a policy consists of $\psi, \lambda$ as well as $(G'_{s'})_{s'=1\dots S}$ that is consistent with the requirements above, i.e.

$$G^0 = r(x, a, s) + \beta E_s G^{0\prime}, \dots, G^I = g^I(x, a, s) + \beta E_s G^{I\prime} \geq \bar{g}_1^I.$$

In order to tie this in with our results in the previous sections, it is useful, however, to choose a slightly different approach. This appears more complicated at first, but will simplify things greatly once we make additional assumptions. We denote by $\widetilde{G}(\gamma, x, s)$ all utilities that can be obtained at a state $(\gamma, x, s)$ without using randomization in the current period, i.e. $G(\gamma, x, s) = \text{conv}(\widetilde{G}(\gamma, x, s))$ and each $G \in \widetilde{G}(\gamma, x, s)$ is associated with a degenerate probability distribution over current actions. A non-standard way of deriving policies now is that a policy specifies $(a, \lambda) \in A(\gamma, x, s)$ as well as $G'_{s'k}$ and $\pi_{s',k}$, $s' = 1, ..., S$ and $k = 1, ..., I$

[3]Formally we now enlarge the state-space to contain agents' utilities, however, when we refer to the state, we mean the current multiplier, current physical state and current shock.

with the property that each $G'_{s'k} \in \widetilde{G}(\gamma + \lambda, x', s')$ and that

$$
\begin{aligned}
G^0 &= r(x,a,s) + \beta E_s \sum_k \pi_{s'k} G'^0_{s'k} \\
G^1 &= g^1(x,a,s) + \beta E_s \sum_k \pi_{s'k} G^{1\prime}_{s'k} \geq \bar{g}^1_1 \\
&\vdots \\
G^I &= g^I(x,a,s) + \beta E_s \sum_k \pi_{s'k} G^{I\prime}_{s'k} \geq \bar{g}^I_1.
\end{aligned}
$$

In other words, the policy today is given by a non-randomized action today and a randomization over all actions tomorrow, shock by shock. By Caratheodory's theorem one only has to randomize over at most $I$ actions in each state tomorrow. In order to make clear that we randomize over a fixed set of continuation utilities, we now use $\pi$ to denote the probabilities of these randomization, instead of $\psi$, the randomization over actions which are obviously implied by this.

## 4.2 A special case

A special case obtains when the sets $G(\gamma, x, s)$ are simple in the sense that they can be written as the convex hull of a finite number of points. In this case, we do not have to introduce an additional continuous state variable.

To illustrate this, we now make the following strong assumptions on the solution set of the problem.

**Assumption 2** *For each admissible $\gamma, x \in \mathcal{X}, x \in \mathcal{S}$, the set $G(\gamma, x, s)$ is a convex polytope, i.e. the convex hull of a finite set of points.*

Note that for the case $I = 1$, i.e. the case of one agent and one principal, the assumptions always hold since $G(\gamma, x, s)$ is simply a closed interval. We consider this case in some detail in the next subsection. For more than one agent, this obviously depends on the formulation of the problem. Even if the set of actions $\mathcal{A}$ is finite, it is not guaranteed that $G(\gamma, x, s)$ will be a polytope – it is quite likely though. However, if the true solution set is not a polytope, the most obvious approach to approximate it numerically is to use a polytope. Furthermore, we will argue below that while solving the problem, one can verify whether or not the true solution set, $G$ is a polytope. It is clear that with this assumption, the state does not has to be enlarged by an infinite set but simply by a finite set of discrete states.

For a given set of points, $P$, whose convex hull forms a polytope, let $\mathcal{V}(P)$ denote the (finite set) of extreme points (i.e. vertices) of this polytope, i.e. the smallest set of points for which $\text{conv}(P) = \text{conv}(\mathcal{V}(P))$. Furthermore, let $g(x,a,s) = (g^0(x,a,s), ..., g^I(x,a,s))$ and let $\mathcal{F} = \{(G_0, ..., G_I) : G_1 \geq \bar{g}^1_1, ..., G_I \geq \bar{g}^I_1\}$ be the set of utilities satisfying the incentive constraint (i.e. feasible utilities). We define a correspondence $\mathcal{G}$, mapping the current state to a finite set of points, $\mathcal{G} : [0, \infty)^I \times \mathcal{X} \times \mathcal{S} \rightrightarrows \mathbb{R}^{I+1}$, by requiring that it satisfies the following functional equation equations

$$
\begin{aligned}
&\mathcal{G}(\gamma, x, s) = & (12)\\
&\mathcal{V}\left[\bigcup_{(a,\lambda)\in A(\gamma,x,s)} \left(\mathcal{F} \cap \operatorname{conv}(\{g(x,a,s)\} + \beta \sum_{s'} \Pi(s,s')\mathcal{G}(\gamma+\lambda, x', s'))\right)\right], \\
&x' = \zeta(x, a, s).
\end{aligned}
$$

In this equation the sum of sets is the Minkowski sum with $\alpha A + \beta B = \{\alpha a + \beta b : a \in A, b \in B\}$.

A solution exists precisely when $G(\gamma, x, s)$ is a polytope – in this case $\mathcal{G}(\gamma, x, s)$ is simply the set of its extreme points. The advantage of writing it like this (as opposed to the system that defined $G(\gamma, x, s)$ in the previous subsection) is that it becomes evident, that these sets can be constructed relatively easily from the knowledge of the correspondence $A(.)$.

When a solution to (12) exists an optimal policy can be constructed as follows. For each $v \in \mathcal{G}(\gamma, x, s)$, define

$$
\begin{aligned}
&\bar{A}(\gamma, x, s, v) = \{(a, \lambda) \in A(\gamma, x, s), (\pi_1, ..., \pi_S) : \\
&v = g(x, a, s) + \beta \sum_{s'} \Pi(s, s') \sum_{v' \in \mathcal{G}(\gamma+\lambda, x', s')} \psi_{s'}(v')v'\}
\end{aligned}
$$

This gives a quasi-policy function in the following sense. The current state now consists of $\gamma, x, s$ as well as a $v \in \mathcal{G}(\gamma, x, s)$. The policy prescribes the current non-randomized action, the current multiplier, $\lambda$ as well as (a possibly degenerate) randomization over next periods states (which in effect means a randomization over actions in the next period). There is in general no guarantee that $\bar{A}$ is single valued, but each selection will produce an optimal policy.

Note that in general one needs to choose a randomization for each shock $s'$ in the next period, i.e. $\pi$ will differ across shocks next period. We will show below that for the special case of one agent and one principal this is not the case – there the same randomization can be chosen for each shock. It is clear why in this general framework this cannot be expected – there is no guarantee that the number of vertices is the same for each shock next periods, therefore it is obviously impossible to find a randomization that does not depend on the shock.

Note that there is only randomization if today at least one incentive constraint is binding.

**Theorem 4** *There is an optimal solution $(a, \lambda, \pi)$ with each $\pi_s$ being a degenerate probability distribution if $(v_1, ..., v_I) \gg (\bar{g}_1^1, ..., \bar{g}_1^I)$.*

To prove the result, observe that in order for $v$ to be an element of

$$
\mathcal{V}\left[\cup_{(a,\lambda)\in A(\gamma,x,s)} \left(\{g(x,s,a)\} + \beta \sum_{s'} \Pi(s,s')\mathcal{G}(\gamma+\lambda, x', s')\right)\right]
$$

it has to be generated by an action next period that is not randomized – each vertex of the sum of polytopes must be the sum of vertices (see e.g. Fukuda (2004)). The result does not seem completely obvious since one might have thought that it could be optimal to randomize in two periods if a IC constraint is only binding this period.

From a computational aspect, polytopes are easy to deal with. There is a sizable literature on polyhedral computation. For example, Fukuda (2004) discusses an algorithm for the Minkowski addition of polytopes that can be applied to very large scale problems. For solving problems with large $I$, the key problem lies in solving for the correspondence $A$, if the true solution set of utilities is a polytope, the computation of $\mathcal{G}(.)$ and $\bar{A}$ is quite straightforward. If one is to solve the functional equation (12) iteratively, one can easily see if the number of extreme points grows with each iteration or if eventually this stabilizes.

## 4.3 A principal and one agent

If there is only one agent (in additional to the principal), Assumption 2 is trivially satisfied and we can set up the problem somewhat more intuitively. Since $G(\gamma, x, s)$ is at most an interval the sets $\mathcal{G}(\gamma, x, s)$ contain at most two points. We can characterize these two points as the one that maximizes the agent's utility and then one that minimizes the agents utility and denote them by $G^H(\gamma, x, s)$ and $G^L(\gamma, x, s)$. Furthermore the probabilities over next periods' continuation payoffs (and hence next periods actions) can now be chosen independently of the shock.

We can consider the following functional equation.

$$
\begin{aligned}
&G^L(\gamma, x, s) = \\
&\min_{(\underline{a},\underline{\lambda})\in A(\gamma,x,s),\pi\in[0,1]} g(x, \underline{a}, s) + \beta E_s \left(\pi G^L(\gamma + \underline{\lambda}, \underline{x}', s') + (1-\pi) G^H(\gamma + \underline{\lambda}, \underline{x}', s')\right) \\
&G^H(\gamma, x, s) = \\
&\max_{(\bar{a},\bar{\lambda})\in A(\gamma,x,s),\pi\in[0,1]} g(x, \bar{a}, s) + \beta E \left(\pi G^L(\gamma + \bar{\lambda}, \bar{x}', s') + (1-\pi) G^H(\gamma + \bar{\lambda}, \bar{x}', s')\right) \\
&\text{ subject to} \\
&\underline{x}' = \zeta(x, \underline{a}, s'),\ \bar{x}' = \zeta(x, \bar{a}, s'), \quad G^H(\gamma, x, s) \geq \bar{g}_1,\ G^L(\gamma, x, s) \geq \bar{g}_1.
\end{aligned}
$$

Given a solution $G^{H*}, G^{L*}$ to this functional equation we define

$$
\begin{aligned}
&\omega_L^*(\gamma, x, s) = \\
&\arg \min_{(a,\lambda)\in A(\gamma,x,s),\pi\in[0,1]} g(x, a, s) + \beta E \left(\pi G^{L*}(\gamma + \lambda, x', s') + (1-\pi) G^{H*}(\gamma + \lambda, x', s')\right) \\
&\text{ subject to } \ x' = \zeta(x, a, s).
\end{aligned}
$$

and

$$
\begin{aligned}
&\omega_H^*(\gamma, x, s) = \\
&\arg \max_{(a,\lambda)\in A(\gamma,x,s),\pi\in[0,1]} g(x, a, s) + \beta E \left(\pi G^{L*}(\gamma + \lambda, x', s') + (1-\pi) G^{H*}(\gamma + \lambda, x', s')\right) \\
&\text{ subject to } \ x' = \zeta(x, a, s).
\end{aligned}
$$

Note that $\omega_L^*$ and $\omega_H^*$ are not policy functions in the classical sense. The policy should specify a randomization $\psi$ and a transition $\lambda$ as a function of the current state. But we have seen that this is not possible. Instead we change the timing and the state. The state now consists of $(\gamma, x, s)$ as well as an element of $\{H, L\}$. The timing now says that given the state, we choose a deterministic action, but also the randomization across actions next period. Note that this randomization will not be conditional on the realized shock next period - but this does not matter since the (IC) holds today.

By construction, these policies are feasible. The mapping could still be set valued, but then we can simply pick any point in that set. Since the solution is feasible, the argument from Theorem 3 can be applied to show that it must also be optimal, i.e. an optimal solution to the original problem can be generated from $\omega$, given an initial $\gamma_0$ that ensures that the (PC) holds.

Finally, note that in the solution to these functional equations $\omega_H^*(\gamma, x, s)$ will trivially set $\pi = 0$ so as to put the maximum probability onto the high continuation payoff for the agent. In the solution to $\omega_L^*(\gamma, x, s)$, it will be optimal to set $\pi = 1$ so long as this does not lead to a violation of the incentive constraint that $G^L(\gamma, x, s) \geq \bar{g}_1$. If this constraint is violated at $\pi = 1$, then it may be optimal to lower $\pi$ since this allows the solution to satisfy the incentive constraint at no additional cost to the overall value of the Lagrangian.

### 4.3.1 The optimal contract

Given these constructions, we can now characterize the dynamics of the optimal contract in terms of $(\gamma, x, s)$ and an element of $\{H, L\}$.

In the initial period the initial physical state is $x_0$, and the promised utility $\bar{g}_2$ associated with the participation constraint must be satisfied. To do so, we select $\gamma_0$ by the requirement that

$$E_s G^H(\gamma_0, x, s) \geq \bar{g}_2 \geq E_s G^L(\gamma_0, x, s),$$

and we select the initial lottery probabilities over $\{H, L\}$, $\zeta$ and $1 - \zeta$, by the requirement that

$$\bar{g}_2 = \zeta E_s G^H(\gamma_0, x, s) + (1 - \zeta) E_s G^L(\gamma_0, x, s).$$

This then determines the full initial state $(\gamma, x, s)$ and an element of $\{H, L\}$. Each period thereafter begins with this initial state.

Within each period thereafter the contract evolves as follows:

1. The element of $\{H, L\}$ determines which policy function $\omega_H^*$ and $\omega_L^*$ is used within the period.

2. The policy function $\omega_i$ determines a continuation triplet $(\gamma + \lambda, x', s')$ for next period and a probability $\pi$ for the lottery over elements of $\{H, L\}$. If the neither the participation or incentive constraint bind, then $\pi$ is simply 0 if the element is $H$ and 1 if the element is $L$. The participation constraint cannot bind if the element is $H$ and if it binds when the element is $L$ then $\pi$ will be less than one. If the incentive constraint binds, then $\pi$ is chosen so that

$$g(x, a_i, s) + \beta E_s \left( \pi G^L(\gamma + \underline{\lambda}, \underline{x}', s') + (1 - \pi) G^H(\gamma + \underline{\lambda}, \underline{x}', s') \right) = \bar{g}_1.$$

where $a_i$ is the minimizing element of the solution set for actions if $i = L$ and maximizing if $i = H$.

3. This lottery is undertaken at the end of the period and an element is selected. This determines the state tomorrow as $(\gamma+\lambda, x', s')$ and the element $a \in \{H, L\}$.

The key thing to note here is that the determination of the lottery probability $\pi$ is very mechanical and does not require solving an ex ante simultaneous equation problem such as would be the case under the utility approach, which we discuss next.

### 4.4 Relation to 'promised utility' approach

It is useful to connect the multiplier approach to the standard 'promised utility method', which was originally developed by Spear and Srivastava (1987), and where the state consists of the physical states as well as of an additional endogenous state variable, promised utility, $G$. This section is somewhat less formal. We focus on the case $I = 1$ and we assume that the promised utility approach has a solution and that the resulting value function is differentiable. With this, we can demonstrate that the value of the Pareto-problem must be identical to the value of the recursive problem of the previous section.

It is useful to change the timing and write value of the Pareto problem as a function of $(G, x)$ before the current shock $s$ is realized. In a Markov setup, we then need to carry around $s_-$ as a state variable and we can write our Pareto problem recursively using the promised utility approach as

$$\begin{aligned}
&V(G,x,s_-) = \max_{(\psi_s)\in\Psi^S,(G'_{s,a})} \min_{\gamma\geq 0} \sum_{s\in\mathcal{S}} \Pi(s_-,s) \sum_{a\in\mathrm{supp}(\psi_s)} \psi_s(a) \\
&\big(r(x,a,s)+\beta V(G'_{s,a},x'_s,s)+\gamma(g(x,a,s)+\beta G'_{s,a}-G)\big) \\
&\text{ subject to} \\
&x'_s = \zeta(x,a_s,s) \\
&p(x,a,s)\geq 0, \quad \forall a\in\mathrm{supp}(\psi_s), \\
&g(x,a,s)+\beta G'_{s,a} \geq \bar{g}_1, \quad \forall a\in\mathrm{supp}(\psi_s).
\end{aligned} \tag{13}$$

Note that we have included the promised utility of the agent $G$ through a Lagrangian multiplier formulation where $\gamma$ is the multiplier on this constraint. Obviously, at the optimal solution it must hold that $\gamma(g(x,a,s)+\beta G'_{s,a}-G)=0$.

Assuming differentiability, each $G$ gives us a unique $\gamma(G)$, since by the envelope theorem $V_G(G,x,s_-)=-\gamma$. Unfortunately that the map is not invertible. For each $\gamma\in\Gamma\subset[0,\infty)$, we can define (assuming differentiability)

$$\begin{aligned}
V^L(\gamma,x,s) &= \min\{V(G,x,s): V_G(G,x,s)=-\gamma\}, \\
V^H(\gamma,x,s) &= \max\{V(G,x,s): V_G(G,x,s)=-\gamma\}.
\end{aligned}$$

We can also define the associated payoffs to the agent as

$$\begin{aligned}
G^L(\gamma,x,s) &= \arg\min\{V(G,x,s): V_G(G,x,s)=-\gamma\}, \\
G^H(\gamma,x,s) &= \arg\max\{V(G,x,s): V_G(G,x,s)=-\gamma\}.
\end{aligned}$$

Note that that we must have that for all $\gamma, x, s$,

$$V^L(\gamma, x, s) + \gamma G^L(\gamma, x, s) = V^H(\gamma, x, s) + \gamma G^H(\gamma, x, s). \tag{14}$$

Given these functions, we can rewrite our recursive Pareto problem in terms of these functions, using the insight from before that the derivative of $V$ in the next period must be equal to $\gamma$ plus the value of the Lagrange parameter attached to this period's (IC). Writing $\gamma$ to denote the optimal $\gamma$ associated with $G$ we obtain

$$\begin{aligned} V(G, x, s_-) = & \max_{(a_s),(\pi_s)} \min_{(\lambda_s)} \sum_s \Pi(s_-, s) \\ & r(x_s, a_s, s) + \beta \left(\pi_s V^L(\gamma + \lambda_s, x_s') + (1 - \pi_s) V^H(\gamma + \lambda_s, x_s')\right) + \\ & \gamma(g(x_s, a_s, s) + \beta(\pi_s G^L(\gamma + \lambda_s, x_s') + (1 - \pi_s) G^H(\gamma + \lambda_s, x_s')) - G) + \\ & \lambda_s(g(x_s, a_s, s) + \beta(\pi_s G^L(\gamma + \lambda_s, x_s') + (1 - \pi_s) G^H(\gamma + \lambda_s, x_s')) - \bar{g}_1) \\ & \text{ subject to } x_s' = \zeta(x_s, a_s, s), \ p(x_s, a_s, s) \geq 0, \ \forall s \end{aligned} \tag{15}$$

Note that the value $V(G, x, s)$ is linear in $G$ for all $\tilde{G}$ that give the same value of $\gamma$ – i.e. we have that

$$V(\tilde{G}, x, s_-) + \gamma \tilde{G} = \text{ constant },$$

whenever $V_G(\tilde{G}, x, s_-) = -\gamma$.

Moreover, if we define $SP(\gamma(G), x, s_-) = V(G, x, S_-) + \gamma(G) G$, we obtain

$$\begin{aligned} SP(\gamma, x, s_-) = & \max_{(a_s),(\pi_s)} \min_{(\lambda_s)} \sum_s \Pi(s_-, s) \\ & r(x_s, a_s, s) + \beta \pi_s \left(V^L(\gamma + \lambda_s, x_s') + (\gamma + \lambda_s) G^L(\gamma + \lambda_s, x_s')\right) \\ & + \beta(1 - \pi_s) \left(V^H(\gamma + \lambda_s, x_s') + (\gamma + \lambda_s) G^H(\gamma + \lambda_s, x_s')\right) + \\ & \gamma g(x_s, a_s, s) + \lambda_s(g(x_s, a_s, s) - \bar{g}_1) \\ & \text{ subject to } x_s' = \zeta(x_s, a_s, s), \ p(x_s, a_s, s) \geq 0, \ \forall s \end{aligned} \tag{16}$$

But note that with Equation (14) the optimal value of this problem can be derived from the optimal value the social planning problem from the previous section. Simply note that

$$SP(\gamma + \lambda_s, x, s) = V^L(\gamma + \lambda_s, x_s') + (\gamma + \lambda_s) G^L(\gamma + \lambda_s) = V^H(\gamma + \lambda_s, x_s') + (\gamma + \lambda_s) G^H(\gamma + \lambda_s, x_s').$$

By changing the timing in the problem of Section 3, it is now clear that we obtain the same value.

# 5 Three simple examples

Here we return to our simple partnership model in which there is a principal and an agent, and the principal must have the participation of the agent in order to run a project which produces a pie of size 1 every period that can be split between the two of them. Each period the principal gets to choose the amount of the pie he eats, $a$, and this implies the amount that the agent eats, $1-a$. The within period reward functions for the principal is $\log(a)$ and for the agent is $\log(1-a)$. Both are expected utility maximizers and both discount the future at rate $\beta$.

The agent has an initial outside opportunity, and then each period draws an outside opportunity. These opportunities put an initial ex ante lower bound on his payoff, which we call the participation constraint, and a conditional lower bound, which we can an incentive constraint. The ex ante outside opportunity which goes into the participation constraint is $\bar{g}_2 = \frac{1}{1-\beta}\log(1/5)$. Assume that in every period there are two states of the world $s \in \{l, h\}$ which are i.i.d. and equi-probable. The conditional opportunity of the agent depends on the shock, $\bar{g}_1(l) = \frac{1}{1-\beta}\log(\epsilon)$ and $\bar{g}_1(h) = \frac{1}{1-\beta}\log(2/3)$. These opportunities determine the conditional incentive constraint.

We will assume that $a$ is bounded between $\varepsilon$ and $1-\varepsilon$, where $\varepsilon$ is a small number that serves to bound the payoffs of the agent and the principal. We will consider two cases: (i) the set of actions $a$ is convex and equal to $[\varepsilon, 1-\varepsilon]$, and (ii) the set of actions is not convex. In the nonconvex case, we will allow for public randomization in order to convexify the set of payoffs. If the set of possible actions $a$ by the principal is convex, then one can show that the Pareto frontier is strictly concave, however if the action set is not convex, then one can easily construct examples in which the Pareto frontier is weakly concave but not strictly so.

## 5.1 Example 1: No flat spots

Assume that the set of possible actions are $a \in [\epsilon, 1-\epsilon]$ for some small $\epsilon > 0$.

The planning problem is given by

$$\max_{a_t} E_{-1} \sum_{t=0}^{\infty} \beta^t \log(a_t)$$

subject to

$$E_{-1} \sum_{t=0}^{\infty} \beta^t \log(1-a_t) \geq \frac{1}{1-\beta}\log(1/5)$$

$$\log(1-a_{h,t}) + E_t \sum_{i=1}^{\infty} \beta^i \log(1-a_{t+i}) \geq \frac{1}{1-\beta}\log(2/3)$$

$$\log(1-a_{l,t}) + E_t \sum_{i=1}^{\infty} \beta^i \log(1-a_{t+i}) \geq \frac{1}{1-\beta}\log(\varepsilon).$$

Clearly the IC cannot bind in state $l$.

This problem is strictly concave, so the first-order conditions are necessary and sufficient. If we denote the multiplier on the PC by $\gamma_0$ and the multipliers on the IC constraint in history state $(h^{t-1}, h)$ and $(h^{t-1}, l)$ by $\beta^t \lambda(h^t)\Pr(h^t)$, we get the first-order condition for $a(h^t)$,

$$\frac{1}{a(h^t)} - \left[\gamma_0 + \sum_{h^j \preceq h^t} \lambda(h^j)\right] \frac{1}{1-a(h^t)} = 0,$$

where "$\preceq$" means predecessor history and includes the current history. The fact that the IC multipliers for each state along an event tree come in additively motivated Marcet and Marimon to focus on the recursive multiplier

$$\gamma(h^{t-1}) = \gamma_0 + \sum_{h^j \preceq h^{t-1}} \lambda(h^j),$$

in which case multiplier in $h^t$ become $\gamma(h^{t-1}) + \lambda(h^t)$ in the above expression. Since this alternative representation is equivalent to the original representation, and since that representation was necessary and sufficient, it follows that this representation is as well. Finally, note here that if we treat $\bar{g}_2$ as a parameter in the ex ante PC, we can solve out for the Pareto frontier by solving for the principal's payoff for each feasible value of $\bar{g}_2 = G$. The resulting Pareto frontier, $P(G)$, which gives the principal's payoff conditional $G$, will be strictly concave.

If we go a step further and consider a recursive planning-problem in which the state variable for the agent is his recursive multiplier $\gamma$, we get Marcet and Marimon's formulation

$$F(\gamma, s) = \max_{a \in A} \min_{\lambda \geq 0} \log(a) + \gamma \log(1-a) + \lambda \left[\log(1-a) - \bar{g}_1(s)\right] + \beta E_s F(\gamma + \lambda, \tilde{s}).$$

The first order conditions with respect to $a$ is simply

$$\frac{1}{a} - \gamma \frac{1}{1-a} - \lambda \frac{1}{1-a} = 0$$

The first order condition with respect to $\lambda$ is

$$\log(1-a) - \bar{g}_1(s) + \beta E_s F' = 0.$$

We obtain directly from these conditions that $a = \frac{1}{1+\gamma+\lambda}$, if the resulting $a \in [\epsilon, 1-\epsilon]$, which will always be the case below. So we only need to solve for $\lambda(\gamma, s)$ and as well as the period zero $\gamma_0$ which is needed to satisfy the participation constraint.

Working recursively at $\bar{\gamma} = 2$, we want to verify that $\lambda(\bar{\gamma}) = 0$, independently of the shock. Note that we can write the utility the IC constraint promises the agent in $h$ as $\bar{g}_1(h) = \frac{1}{1-\beta} \log(\frac{\bar{\gamma}}{1+\bar{\gamma}})$. Furthermore, under the assumption that $\lambda = 0$, the value-function is given by

$$F_l(\bar{\gamma}) = F_h(\bar{\gamma}) = \frac{1}{1-\beta} \left( \log(\frac{1}{1+\bar{\gamma}}) + \bar{\gamma} \log(\frac{\bar{\gamma}}{1+\bar{\gamma}}) \right).$$

In shock $h$, the first order condition with respect to $\lambda$ at $\lambda = 0$ becomes

$$\log(\frac{\gamma}{1+\gamma}) - \frac{1}{1-\beta} \log(\frac{\bar{\gamma}}{1-\bar{\gamma}}) + \frac{\beta}{1-\beta} (\log(\bar{\gamma}) - \log(1+\bar{\gamma})) = 0,$$

so in fact $\lambda = 0$ is the optimal policy if $\bar{\gamma} = 2$. In shock $l$ it is easy to see that the $\lambda \geq 0$ constraint becomes binding and so $\lambda = 0$ is also the optimal policy.

Working backwards, for any $\gamma < \bar{\gamma}$ it is easy to verify, by the above argument, that in shock $h$ the optimal policy is given by $\lambda(\gamma) = \bar{\gamma} - \gamma$ and so $a = \frac{1}{1+\bar{\gamma}}$. The value function is then

$$\begin{aligned} F_h(\gamma) &= \frac{1}{1-\beta}\left(\log(\frac{1}{1+\bar{\gamma}}) + \bar{\gamma}\log(\frac{\bar{\gamma}}{1+\bar{\gamma}})\right) - (\bar{\gamma} - \gamma)\bar{g}_2 \\ &= \frac{1}{1-\beta}\left(\log(\frac{1}{1+\bar{\gamma}}) + \gamma\log(\frac{\bar{\gamma}}{1+\bar{\gamma}})\right).^4 \end{aligned}$$

In shock $l$, we claim that $\lambda = 0$ is the optimal policy, whenever $\gamma \geq \frac{1}{1-\epsilon} - 1$. If $\lambda = 0$,

$$F_l(\gamma) = \log(\frac{1}{1+\gamma}) + \gamma\log(\frac{\gamma}{1+\gamma}) + \frac{\beta}{2}\left[F_l(\gamma) + F_h(\gamma)\right],$$

and substituting for $F_h(\gamma)$ and $\bar{\gamma}$, the value function is given by

$$F_l(\gamma) = \frac{1}{1-\beta/2}\left(\log(\frac{1}{1+\gamma}) + \gamma\log(\frac{\gamma}{1+\gamma}) + \frac{\beta}{2-2\beta}(\log(1/3) + \gamma\log(2/3))\right).$$

One can easily check that the first order conditions with respect to $\lambda$ give a corner solution $\lambda = 0$.

Finally the initial $\gamma_0$ is now pinned down by the participation constraint through the value function of the agent.[5]

## 5.2 Example 2: Flat spots

We now change the previous example and assume that $a \in \{\epsilon\} \cup [0.5, 1-\epsilon]$. There is now a potential flat spot in the Pareto-frontier at

$$\gamma^* = \frac{\log(0.5) - \log(\epsilon)}{\log(1-\epsilon) - \log(0.5)} > 1.$$

For an $\gamma > \gamma^*$ the corner solution $a = \epsilon$ is optimal, and the constraint that $\lambda \geq 0$ binds, hence $\lambda = 0$, the (IC) constraint does not bind and $F_s(\gamma) = \frac{1}{1-\beta}(\log(\epsilon) + \gamma\log(1-\epsilon))$ for both $s = 1, 2$. At $\gamma = \gamma^*$ the optimal solutions (to the recursive Pareto-problem) are all randomizations over 0.5 and $\epsilon$, yielding 0.5 or $1-\epsilon$ to the agent. Since any $\gamma > \gamma^*$ implies the agent getting $1-\varepsilon$ forever, the IC constraints cannot possibly bind at $\gamma^*$ and $\lambda = 0$. Note that if $\beta$ is so small that

$$\log(0.5) - \frac{1}{1-\beta}\log(2/3) + \frac{\beta}{1-\beta}\log(1-\epsilon) < 0,$$

it will never be optimal to put positive weight on 0.5 in the $h$ state and randomization only occurs in the $l$ state. Clearly not all these randomizations are optimal solutions to the principal agent problem (even if they only occur in the $l$-state).

[5] Marcet and Marimon don't include an initial participation constraint and hence in their formulation, $\gamma_0 = 0$.

If we let $G_s^H(\gamma^*)$ denote the high payoff to the agent in state $s$ and $G_s^L(\gamma^*)$ the low, then randomizing over extreme payoffs in the appropriate manner will allow us to achieve the efficient outcome. Clearly at $\gamma = \gamma^*$, the highest possible outcome for the agent is to set his consumption equal to $1-\epsilon$ forever, and this implies that $G_s^H(\gamma^*) = \frac{1}{1-\beta}\log(1-\epsilon)$, $s = l, h$, and that the ex ante high payoff $G^H(\gamma^*)$ is equal to this as well.

Next consider recursively defining $G_s^L(\gamma^*)$ as

$$G_s^L(\gamma^*) = \min_{\pi\in[0,1], a\in\{0.5, 1-\epsilon\}} \log(1-a) + \beta\left(\pi G^H(\gamma^*) + (1-\pi)G^L(\gamma^*)\right)$$

subject to

$$G_s^L(\gamma^*) \geq \bar{g}_1(s)$$

and where

$$G^L(\gamma^*) = .5\left[G_l^L(\gamma^*) + G_h^L(\gamma^*)\right].$$

Clearly, we will want to make $\pi$ and $a$ as small as possible subject to the incentive constraint. Since the incentive constraint has to hold ex-post, we cannot randomize over $a$ to satisfy this constraint. However, we can do so by randomizing over $G^L$ and $G^H$, which means that we are randomizing over actions tomorrow. Since the incentive constraint doesn't bind in the low state, it follows that $\pi = 0$, $a = 0.5$, and

$$G_l^L(\gamma^*) = \log(0.5) + \frac{\beta}{2}\left[G_l^L(\gamma^*) + G_h^L(\gamma^*)\right]. \tag{17}$$

Since the incentive constraint binds in the high state $G_h^L(\gamma^*) = \bar{g}_1(h)$. Equation (17) then yields directly

$$G_l^L(\gamma^*) = \frac{1}{1-\beta/2}(\log(0.5) + 0.5\beta\bar{g}_1(h)).$$

The associated actions are clear in the low state – in the high state they depend on $\epsilon$ and the discount factor. If there exists a $\pi$ such that

$$\log(0.5) + \beta\left(\pi(0.5G_h^H + 0.5G_l^H) + (1-\pi)(0.5G_h^L + 0.5G_l^L)\right) = \bar{g}_1(h),$$

then the optimal action is 0.5, otherwise it must be $\epsilon$. Note that there might be a region where the action is not unique and both 1/2 and $\epsilon$ to the job with the right probabilities.

For any $\gamma_0 < \gamma^*$ in shock $h$ the (IC) constraint binds, the optimal $\lambda(\gamma_0) = \gamma^* - \gamma_0$ and the associated action is $\omega_L^H(h)$ together with a randomization over next periods' actions that gives $\bar{g}_1(h)$ for the agent. In shock $l$, on the other hand, as in the example above, if $\gamma_0 \geq \frac{1}{1-\epsilon} - 1$ the optimal $\lambda = 0$ and the unique policy is determined by $a = \frac{1}{1+\gamma_0}$. The participation constraint pins down $\gamma_0$. The reason why one does not have to keep track of $G^H$ and $G^L$ for $\gamma < \gamma^*$ is simple: At $h$, there is initially no randomization, the $\gamma + \lambda$ brings one to the flat spot but it is clear that the strategy is uniquely determined as the one that yields $\bar{g}_1(h)$.

Note that the randomization $\pi$ is identical for both shocks in the subsequent period. This is only possible because we randomize over the two extreme points of two intervals, describing the sets $\mathcal{G}(s')$ for the two possible shocks $s' = 1, 2$ in the next period.

## 5.3 Example 3: How do Flat Spots Propagate

We now change the example by assuming that the action space can stochastically switch from the continuous action space $(c)$ to the discrete action space $(d)$ and stay there forever. The state space is now $s \in \{(l,c),(h,c),(l,d),(h,d)\}$, where the draws of $l$ and $h$ continue to be i.i.d. and equiprobable, and the probability of switching from $c$ to $d$ next period is $\rho$, while $d$ is an absorbing state. If the state is $(i,c)$ for $= l,h$ then the action space is $A(i,c) = [\epsilon, 1-\epsilon]$, while if it is $(i,d)$ it is $A(i,d) = \{\epsilon\} \cup [0.5, 1-\epsilon]$. We assume that the initial states are either $(l,c)$ or $(l,d)$.

The recursive planning problem is

$$F(\gamma, s) = \max_{a \in A(s)} \min_{\lambda \geq 0} \log(a) + \gamma \log(1-a) + \lambda\left[\log(1-a) - \bar{g}_1(s)\right] + \beta E_s F(\gamma+\lambda, \tilde{s}),$$

and the first-order conditions are still

$$\frac{1}{a} - \gamma\frac{1}{1-a} - \lambda\frac{1}{1-a} = 0$$

with respect to $a$ and

$$\log(1-a) - \bar{g}_1(s) + \beta E_s F' = 0$$

with respect to $\lambda$, which implies that $a = \frac{1}{1+\gamma+\lambda}$, if the action space is $A(i,c)$ and the resulting $a \in [\epsilon, 1-\epsilon]$.

Clearly, when the action space is $A(i,d)$, the solution to this problem will work just like in the example 2, and there will be a flat spot on the conditional Pareto frontier at slope $\gamma^*$. For this reason, continue to let $G_l^j$ for $j = H$ and $L$ denote the high and low payoffs respectively in the discrete case when $s = l$. In state $(l,c)$, the payoff to the agent, $W_l(\gamma)$, will be implicitly given by

$$\begin{aligned} W_l(\gamma) &= \log(1-a(\gamma+\lambda)) + \frac{\beta}{2}(1-\rho)W_l(\gamma+\lambda) \\ &\quad +\rho\frac{\beta}{2}\left[\left(\pi G_l^H(\gamma+\lambda) + (1-\pi)G_l^L(\gamma+\lambda)\right)\right] \\ &\quad +\frac{\beta}{2}\left[(1-\rho)+\rho\right]\bar{g}_1(h), \end{aligned}$$

where $\lambda$ is chosen to satisfy the incentive constraint that this payoff must be at least $\bar{g}_1(l)$. If $\lambda > 0$ then this payoff must be $\bar{g}_1(l)$, and hence is uniquely defined. If $\gamma \neq \gamma^*$ and $\lambda = 0$, then $G_l^H(\gamma) = G_l^L(\gamma)$ and the payoff is again uniquely defined. However if $\gamma = \gamma^*$ and $\lambda = 0$, then the $G_l^H(\gamma) > G_l^L(\gamma)$ and the ability to randomize over this two continuation payoffs will induce a flat section on the Pareto frontier today.

Similarly, in state $(h,c)$ the payoff to the agent, $W_h(\gamma)$ will be implicitly given by

$$\begin{aligned} W_h(\gamma) &= \log(1-a(\gamma+\lambda)) + \frac{\beta}{2}(1-\rho)W_h(\gamma+\lambda) \\ &\quad +\rho\frac{\beta}{2}\left[\left(\pi G_l^H(\gamma+\lambda) + (1-\pi)G_l^L(\gamma+\lambda)\right)\right] \\ &\quad +\frac{\beta}{2}\left[(1-\rho)+\rho\right]\bar{g}_1(h), \end{aligned}$$

where $\lambda$ is chosen to satisfy the incentive constraint that this payoff must be at least $\bar{g}_1(h)$. If with $\gamma = \gamma^*$, $\lambda = 0$ and $\pi = 1$ the payoff $W_h(\gamma) > \bar{g}_1(h)$, then here too the payoff to the agent will be indeterminate when $\gamma = \gamma^*$.

This example illustrates that flat sections in the continuation Pareto frontier do induce flat spots in the current Pareto frontier, but only at the specific value of the multiplier.[6]

# 6 Computational Considerations

We show in this section that the insights from the simple example above carry over to more complicated models. For simplicity we assume that there is no physical state. For models with a physical state variable (such as for example capital), many of the insights of this section will still hold true as long as one makes strong enough assumptions on the law of motion, $\zeta$.

We propose a simple value function iteration to solve for the optimal policy even when there are flat spots in the Pareto-frontier. In the case without physical state, the recursive planner problem simplifies to

$$\begin{aligned}
&F(\gamma, s) = \max_{\psi \in \Psi} \min_{\lambda \geq 0} \sum_{a \in \operatorname{supp}(\psi)} \psi(a) \\
&\left[ r(a, s) + \sum_{i=1}^{I} \left( \gamma^i g^i(a, s) + \lambda_a^i (g^i(a, s) - \bar{g}_{1s}^i) \right) + \beta E_s F(\gamma + \lambda_a, s') \right] \\
&\quad \text{subject to} \\
&p(a, s) \geq 0, \quad \forall a \in \operatorname{supp}(\psi).
\end{aligned} \tag{18}$$

Standard arguments show that for a given $s$, $F$ is continuous in $\gamma$. To see that it is also convex it is useful to consider the non-recursive social planning problem (7).

$$\begin{aligned}
&SP(\gamma_0, s_0) = \max_{(\psi_t) \in \Psi^\infty} E_{s_0}^{\psi} \left[ \sum_{t=0}^{\infty} \beta^t \left( r(a_t, s_t) + \sum_{i=1}^{I} \gamma_i g^i(a_t, s_t) \right) \right] \text{ subject to} \\
&\operatorname{supp}(\psi_t) \subset \mathcal{A}(h^{t-1}, s_t) \text{ for all } h^{t-1}, s_t \\
&g^i(a_t, s_t) + E_t^{\psi} \sum_{n=1}^{\infty} \beta^n g^i(a_{t+n}, s_{t+n}) \geq \bar{g}_1^i \quad \forall a_t \in \operatorname{supp}(\psi_t), \ \forall t, \ \forall i
\end{aligned}$$

We proved above that $SP(\gamma, s) = F(\gamma, s)$ and furthermore it is easy to see that that for any $s \in s$, $\gamma, \tilde{\gamma} \geq 0$ and $\mu \in [0, 1]$, $SP(\mu\gamma + (1 - \mu)\tilde{\gamma}, s) \leq \mu SP(\gamma) + (1 - \mu)SP(\tilde{\gamma})$. For this just note that if $(\psi_t^*)$ solves the social planner problem for $\gamma^* = \mu\gamma + (1 - \mu)\tilde{\gamma}$, it is feasible for

[6] This result is sensitive to the assumption that the principal and the agent discount at the same rate.

both $\gamma$ and $\tilde{\gamma}$. Therefore

$$\mu SP(\gamma, s) + (1-\mu) SP(\tilde{\gamma}, s) \geq$$
$$\mu E_{s_0}^{\psi^*} \sum_{t=0}^{\infty} \beta^t \left( r(a_t, s_t) + \sum_{i=1}^{I} \gamma_i g^i(a_t, s_t) \right) + (1-\mu) E_{s_0}^{\psi^*} \sum_{t=0}^{\infty} \beta^t \left( r(a_t, s_t) + \sum_{i=1}^{I} \tilde{\gamma}_i g^i(a_t, s_t) \right) =$$
$$SP(\mu\gamma + (1-\mu)\tilde{\gamma})$$

Let $\nabla F(\gamma, s)$ denote the subdifferential of $F$ with respect to $\gamma$. This possibly multivalued map is (cyclically) monotonically increasing in $\gamma$. If $F$ is differentiable at$\gamma$, the map is single-valued and equal to the gradient of $F$ (see Rockefellar (1970)). The points where $F$ fails to be differentiable are precisely those that correspond to a flat spot in the Pareto-frontier. It is also standard to see that given an action $a$, an interior solution for $\lambda$ must satisfy the first order condition

$$g(a, s) - \bar{g}_{1s} + \beta E_s \partial(s') = 0, \tag{19}$$

with $\partial(s') \in \nabla F(\gamma + \lambda, s')$ for all $s'$.

It is clear that for a given $(\gamma, s)$ and a given $\lambda$ the set of optimal actions (that are associated with that $\lambda$) can be determined without knowledge of $F$. If the problem is smooth it is simply the set of zeros of the equations

$$D_a r(a, s) + (\gamma + \lambda) Dg(a, s) + \eta Dp(a, s) = 0, \quad \eta_i p_i(a, s) = 0, \tag{20}$$

that satisfy $p(a, s) \geq 0$. In the strictly concave case, this system has a unique solution which can be found with standard methods, otherwise all solutions to this system can be obtained using Gröbner bases if the reward functions are algebraic and the dimension of the action space is not too large (see Kubler and Schmedders (2010)) – this method has the advantage that $(\gamma+\lambda)$ can be treated as a parameter and the Gröbner basis only needs to be computed once.

If the problem is not smooth, as in the example, one needs to look first for solutions in the smooth region and then go through all endpoints (or discrete choices) to enumerate the support of all possible randomizations.

Note that by the first order conditions (19) for a given $\lambda$, if one of the reward functions is strictly monotone in the action, there can only be more than one solution if $\lambda = 0$ or if we are at a point where $F(\gamma + \lambda, s')$ is not differentiable for some $s'$.

Also note that if $a(\gamma + \lambda)$ and $a(\gamma + \tilde{\lambda})$ and solve Equation (20) for $\lambda$ and $\tilde{\lambda}$ respectively, and if $\lambda < \tilde{\lambda}$, we must have $g(a(\gamma + \lambda), s) < g(a(\gamma + \tilde{\lambda}), s)$. Since $\nabla F$ is monotone, the first order condition (19) then implies directly that there cannot be two different $(a, \lambda)$ and $(\tilde{a}, \tilde{\lambda})$ with $\tilde{\lambda} \neq \lambda$ that both solve the Bellman equation (10).

These observations suggest a simple value function iteration to solve the functional Equation (10) and to implement our approach computationally.

We start with an initial guess of a single-valued policy and differentiable and convex value function. As in standard value function iteration (see e.g. Judd (1998)), we iterate backwards, solving the Bellman equation at a finite number of predetermined points in each iteration. This is possible, because we can control for possibility of multiplicity in a simple fashion. We check separately if there are solutions for $\lambda = 0$. If not, we use standard methods to find a solution for an interior $\lambda$. Only if the associated $\lambda > 0$ leads to a nondifferentiability of tomorrow value function we need to again use all solution methods to find all solutions in $a$ for that given $\lambda$. At the same time, we perform value function iteration on the agents' utilities $G(\gamma, s)$ and approximate the optimal policy $\bar{A}(\gamma, s, v)$ For the case $I > 1$ it can be verified at each stage that if there is a flat spot, the set of agents' utility in fact does form a polytope. For the simple case $I = 1$, in the case of a flat spot, we simply need to keep track both of $G^L$ and $G^H$ and include into the policy the optimal randomization over next periods' continuation. Value function iteration can be done parallel to keeping track of the vertices of the set $G(\gamma, s)$.

# 7 Conclusion

The recursive multiplier approach in Marcet and Marimon (1994) can be extended to models with flat spots in the Pareto-frontier. In order for our method to be computationally tractable, the flat spots must be polytopes. This is obviously always the case if there is only one agent in addition to the principal. In the case of several agents, this will naturally be the case if the flat spots are generated lotteries over a finite number of actions. On the other hand, it is clearly possible to construct problems where the flat spot is a general convex set. For problems like this, neither the promised utility approach nor our method is likely to be feasible.

# References

[1] Aiyagari, R., A. Marcet, T. Sargent and J. Seppala (2002), *Journal of Political Economy*, 110(6), 1220-1254.

[2] Atkeson, A. and H. Cole, (2005) A Dynamic Theory of Executive Compensation and Optimal Capital Structure, N.B.E.R. working paper #11083.

[3] Attanasio, O. and J. Rios-Rull (2000), Consumption Smoothing in Island Economies: Can Public Insurance Reduce Welfare?, *European Economic Review*, 44(7), 1225-1258.

[4] Chien, Y., H. Cole and H. Lustig, (2007), A Multiplier Approach to Understanding the Macro Implications of Household Finance, N.B.E.R. working paper #13555.

[5] Cole, H. and E. Prescott (1997), Valuation Equilibrium with Clubs, *Journal of Economic Theory*, 74, 19-39.

[6] Cooley, T., R. Marimon and V. Quadrini (2004), Aggregate Consequences of Limited Contract Enforceability, *Journal of Political Economy*, 112(4), 817-847.

[7] Dechert, W.D., (1982), Lagrange multiplier in infinite horizon discrete time optimal control models, *Journal of Mathematical Economics*, 9, 285-302.

[8] Fukuda, K, (2004) From the zonotope construction to the Minkwoski sum of convex polytopes, *Journal of Symbolic Computation*, 38, 1261-1272.

[9] Judd, K. (1998), *Numerical Methods in Economics*, MIT Press, Cambridge.

[10] Kehoe, P. and F. Perri (2002), International Business Cycles with Endogenous Incomplete Markets, *Econometrica*, 70(3), 907-928.

[11] Khan, A., R. King and A. Wolman (2003), Optimal Monetary Policy, Review of Economic Studies, 70(4), 825-860.

[12] Kubler, F. and K. Schmedders (2010), Tackling Multiplicity of Equilibria with Gröbner Bases, *Operations Research* 58, 1037-1050.

[13] Luenberger, D.G. (1969), *Optimization by vector space methods*, Wiley.

[14] Marcet, A. and R. Marimon, (1992), Communication, Commitment, and Growth, *Journal of Economic Theory*, 58, 219-49.

[15] Marcet, A. and R. Marimon, (1994), Recursive contracts., discussion paper Universitat Pompeu Fabra.

[16] Messner,M. and N. Pavoni, (2004), On the recursive saddle point method, IGIER working paper 255, Universita Bocconi.

[17] Prescott, E. and R. Townsend (1984A), General Competitive Analysis in an Economy with Private Information, *International Economic Review*, 25(1), 1-20.

[18] Prescott, E. and R. Townsend (1984A), Pareto Optima and Competitive Equilibria with Adverse Selection and Moral Hazard, *Econometrica*, 52(1), 21-46.

[19] Rockefellar, R.T., (1970), *Convex Analysis*, Princeton University Press.

[20] Rogerson, R. (1988), Indivisible Labor, Lotteries and Equilibrium, *Journal of Monetary Economics*, 21, 3-16.

[21] Rustichini, A. (1998), Lagrange multiplier in incentive-constrainted problems, *Journal of Mathematical Economics*, 29, 365-380.

[22] Spear, S. and S. Srivastava (1987), On Repeated Moral Hazard with Discounting, *Review of Economic Studies*, 54, 599-617.

[23] Stokey, N., and R. Lucas, with E. Prescott (1989) *Recursive Methods in Economic Dynamics*, Harvard University Press.