RECURSIVE CONTRACTS

Albert Marcet
London School of Economics and IAE (CSIC) - BarcelonaGSE
and
Ramon Marimon
European University Institute and UPF - BarcelonaGSE

This version: **January, 2011**[*]

Abstract

We obtain a recursive formulation for a general class of contracting problems involving incentive constraints. These constraints make the corresponding maximization (*sup*) problems non recursive. Our approach consists of studying a recursive Lagrangian. Under standard general conditions, there is a recursive *saddle point* (*infsup*) functional equation (analogous to Bellman's equation) that characterizes the recursive solution for the planner's problem and the individual values. Our approach has been applied to a large class of dynamic contractual problems; such as: contracts with limited enforcement, optimal policy design with implementability constraints, or dynamic political economy models.

1 Introduction

Recursive methods have become a basic tool for the study of dynamic economic models. For example, Stokey, et al. (1989) and Ljungqvist and Sargent (2000) describe a large number of macroeconomic models that can be analyzed using recursive methods. A main advantage of such approach is that it characterizes optimal decisions -at any time t- as a time-invariant functions of a small set of state variables. However, a key condition in standard dynamic programming techniques is that only past variables can influence the set of feasible current actions. Unfortunately, many interesting economic problems do not satisfy such

[*]This is a substantially revised version of previously circulated papers with the same title (e.g. Marcet and Marimon1998,1999). We would like to thank Fernando Alvarez, Truman Bewley, Edward Green, Robert Lucas, Andreu Mas-Colell, Fabrizio Perri, Edward Prescott, Victor Rios,Thomas Sargent, Robert Townsend and Jan Werner for comments on earlier developments of this work, all graduate students who have struggled through a theory in progress and, in particular, to Matthias Mesner and Nicola Pavoni for pointing out a problem overlooked in previous versions. Support from MCyT-MEyC of Spain and the hospitality of the Federal Reserve Bank of Minneapolis is acknowledged.

condition. For example, in contracting problems where agents are subject to intertemporal participation, or other intertemporal incentive constraints, the future development of the contract determines its feasibility. Similarly, in models of optimal policy design agent's reactions to government policies are taken as constraints and, therefore, future actions limit the set of current feasible actions available to the government.

In this paper we provide an integrated approach for a recursive formulation of a large class of dynamic models with constraints that depend on expectations of functions of future control variables. Even though these models are not recursive in the standard sense, by reformulating them as an equivalent recursive saddle point problem we obtain a recursive formulation.

We build on traditional tools of economic analysis, such as duality theory (in optimization problems), fixed point theory (in infinite dimensional spaces), and dynamic programming. We proceed in three steps. We first study the planners problem with incentive constraints (PP) as an infinite-dimensional maximization problem, for which standard duality theory applies. Second, we show the equivalence between the planner's problem and a modified saddle point problem (SPP). Third, we extend dynamic programming theory to show that the (SPP) has a recursive formulation in the sense that it satisfies a saddle point functional equation (SPFE) which generalizes Bellman's equation.

The resulting saddle point problem (SPP) expands the set of state variables to include new variables that summarize the evolution of the lagrange multipliers of the original (PP) problem. Such transformation creates some technical difficulties since the new (co)state variables can not be bounded. Fortunately, we can exploit the resulting homogeneity properties of the return function and, in this way, we are able to extend the standard contraction mapping approach to establish the relationship between SPP and the SPFE.

We show that solving the lagrangean (SPP) is equivalent to solving the recursive SPFE without concavity assumptions. This is important because incentive constraints may not have a convex structure. If concavity is satisfied, then solving the SPP (and, therefore, the SPFE) is equivalent with solving the maximization problem PP. In the absence of concavity, as in any application of lagrangean theory, our SPFE characterization is sufficient but it may not be necessary for a solution. Our sufficiency results are very general, although we assume that solutions are unique. Nevertheless, as more recently Marimon, Messner and Pavoni (2011) have shown, our approach can be easily generalized to the case where the SPFE has multiple solutions.

As standard recursive methods have proved to be useful to study many dynamic economic models -specially, but not only, in macroeconomics,- our approach has a wide range of applications. It has already proved to be useful to study models such as: growth and business cycles with possible default (Marcet and Marimon (1992), Kehoe and Perri (1998), Cooley, Marimon and Quadrini (2000)), social insurance (Attanasio and Rios-Rull (2000)) and optimal fiscal and monetary policy design with incomplete markets (Rojas (1993), Marcet, Sargent and Seppala (1996)). For brevity, however, we do not present further applications here and limit the presentation of the theory to the case of full information[1]. Our approach is related to other existing approaches that study dynamic models with expectations constraints. In particular, to the pioneer works of Abreu, Pearce and Stacchetti (1990), Green (1987) and Thomas and Worrall (1988) and, among others, the more recent contributions of Kocherlakota (1996) and Rustichini (1998a). We briefly discuss how these, and others, works relate to ours in Section 4, after presenting the main body of the theory in Sections 2 and 3 (while most proofs are contained in the Appendix).

2 Formulating contracts as recursive saddle-point problems

In this section we briefly present our approach and develop the *saddle -point* formulation of problems with intertemporal constraints. We start by considering problems that have the following representation:

$$\textbf{PP} \qquad \sup_{\{a_t\}} \mathrm{E}_0 \sum_{t=0}^{\infty} \beta^t r(x_t,\, a_t,\, s_t) \tag{1}$$

$$\text{s.t. } x_{t+1} = \ell(x_t,\, a_t,\, s_{t+1}), \;\; p(x_t, a_t,\, s_t) \geq 0, \tag{2}$$

$$\mathrm{E}_t \sum_{n=1}^{N_j+1} \beta^n h_0^j(x_{t+n},\, a_{t+n},\, s_{t+n}) + h_1^j(x_t,\, a_t,\, s_t) \geq 0, \tag{3}$$

$$j = 1, ..., l;\;\; t \geq 0, \quad x_0 = x,\, s_0 = s$$

$$a_t \text{ measurable with respect to } (\ldots, s_{t-1}, s_t).$$

Standard dynamic programming methods only consider constraints of the from (**??**) (see, for example, Stokey, *et al.* (1989) and Cooley, (1995)). However, constraints of the form (**??**) are not a special case of (**??**), since they involve expected values of future variables[2]. We know from Kydland and Prescott (1977) that, under these constraints, the usual Bellman equation is not satisfied, the solution is *not,* in general, of the form $a_t = f(x_t, s_t)$ for all t and the whole history of past shocks s_t can matter for today's optimal decision. By letting $N_j = \infty$ **PP** covers a large class of problems where discounted present values enter the implementability constraint. For example, long term contracts with *intertemporal participation constraints* take this form[3] Alternatively, by

[1]Antonio Mele (2010) extends our approach to moral hazard problems.

[2]Notice that expressing (**??**) in the form $v(x_t, s_t) - \psi(x_t, s_t) \geq 0$, where v is the value function of **PP** and ψ some exogenously given participation constraint, is not a special case of (**??**) since v is not known a priori. On the other hand, combining **??**) and (**??**) accounts for a broad class of constraints. For example, a nonlinear participation constraint of the form $g(\mathrm{E}_t \sum_{n=0}^{\infty} \beta^n h(x_{t+n},\, a_{t+n},\, s_{t+n}), x_t, a_t, s_t) \geq 0$ can easily be incorporated in our framework with one constraint of the form (**??**), $g(w_t, x_t, a_t, s_t) \geq 0$ (with control variables (w_t, a_t)), and a constraint of the form (**??**), $\mathrm{E}_t \sum_{n=0}^{\infty} \beta^n h(x_{t+n},\, a_{t+n},\, s_{t+n}) = w_t$.

[3]Examples using related methods can be found in Kocherlakota (1996) and Kletzer and Wright (1998), and using our approach in Marcet and Marimon (1992), Kehoe and Perry (1998), Attanasio and Rios-Rull (2000). Sargent and Ljungqvist (2002) provides examples with different approaches.

letting $N_j = 0$ **PP** covers problems where intertemporal reactions of agents must be taken into account. For example, dynamic Ramsey problems, where the government chooses policy variables subject to intertemporal implementability constraints, have this form[4]. We consider the two canonical cases $N_j = \infty$ and $N_j = 0$, while other intermediate cases can be easily incorporated (without loss of generality, let $N_j = \infty$, for $j = 0, ..., k$, and $N_j = 0$ for $j = k+1, ..., l$).

We consider a more general class of problems parameterized by μ,

$$\mathbf{PP}_\mu \qquad \sup_{\{a_t\}} \mathrm{E}\left[\sum_{j=0}^{l}\sum_{t=0}^{N_j}\beta^t \mu^j h_0^j(x_t,\, a_t,\, s_t) \mid s\right]$$

$$\text{s.t. } x_{t+1} = \ell(x_t,\, a_t,\, s_{t+1}),\;\; p(x_t, a_t,\, s_t) \geq 0, \tag{4}$$

$$\mathrm{E}_t \sum_{n=1}^{N_j+1} \beta^n h_0^j(x_{t+n},\, a_{t+n},\, s_{t+n}) + h_1^j(x_t,\, a_t,\, s_t) \geq 0,\;\; t \geq 0 \tag{5}$$

$$x_0 = x,\, s_0 = s,\;\; N_j = \infty,\;\; j = 0, ..., k, N_j = 0, j = k+1, ...l$$

$$a_t \text{ measurable with respect to } (\ldots, s_{t-1}, s_t).$$

It is easy to see that **PP** is a special case of $\mathbf{PP}_\mu$. It only requires to identify the function h_0^0 with the function r, set $\mu = (1, 0, ..., 0)$ and – provided that r is bounded – choose a h_1^0 for which the corresponding participation constraint is never binding. It should be noticed that the value function, when well defined, takes the form $V_\mu(x,\, s) = \mu v_\mu(x,\, s)$.This Pareto-welfare form will play a role in some of ours characterization results. Notice that $\mathbf{PP}_\mu$ is an infinite dimensional maximization problem which, under relatively standard assumptions, has a solution in state (x, s) which is a *plan*[5] $\boldsymbol{a} \equiv \{a_t\}_{t=0}$, where $a_t(\ldots, s_{t-1}, s_t)$ is a state-contingent action (Proposition 1).

An intermediate step in our approach is to transform program $\mathbf{PP}_\mu$, first, into a saddle-point problem in the following way. Consider writing the Lagrangean for $\mathbf{PP}_\mu$ when a lagrange multiplier $\gamma \in R^{l+1}$ is attached *only* to the forward looking constraints in period $t = 0$ and the remaining constraints are left as constraints. The Lagrangean then is

$$L(\mathbf{a}, \boldsymbol{\gamma}; \mu) \;\; = \;\; \mathrm{E}\left[\sum_{j=0}^{l}\sum_{t=0}^{N_j}\beta^t \mu^j h_0^j(x_t,\, a_t,\, s_t) \mid s\right] + \tag{6}$$

$$\sum_{j=0}^{l}\gamma^j \mathrm{E}\left[\sum_{t=1}^{N_j+1} \beta^t h_0^j(x_t,\, a_t,\, s_t) + h_1^j(x_0,\, a_0,\, s_0) \mid s\right] \tag{7}$$

If we find a saddle point of L subject to (**??**) the usual equivalences between this saddle point and the optimal allocation of $\mathbf{PP}_\mu$ can be exploited. Using simple

[4]Examples using the "primal approach"can be found in Chari *et al.* (1995) and Lucas and Stokey (1983); using related methods in Chang(1998) and Phelan and Stacchetti (1999), and using our approach in Rojas (1993), and previous working paper versions of this paper.

[5]We use the bold notation to denote sequences of measurable functions.

algebra it is easy to show that L can be rewritten as the objective function in the following saddle point problem[6]:

$$
\begin{aligned}
\mathbf{SPP}_{\mu} \qquad & \inf_{\gamma \in R_{+}^{l+1}} \sup_{\{a_t\}} \mu h_0(x_0,\, a_0,\, s_0) + \gamma h_1(x_0,\, a_0,\, s_0) \\
& + \beta \mathrm{E}\left[\sum_{j=0}^{l} \varphi^j(\mu,\, \gamma) \sum_{t=0}^{N_j} \beta^t \; h_0^j(x_{t+1},\, a_{t+1},\, s_{t+1}) \mid s_0\right] \qquad (8) \\
\text{s.t. } & x_{t+1} = \ell(x_t,\, a_t,\, s_{t+1}), \;\; p(x_t, a_t,\, s_t) \geq 0, \quad t \geq 0 \\
& \mathrm{E}_t \sum_{n=1}^{N_j+1} \beta^n h_0^j(x_{t+n},\, a_{t+n},\, s_{t+n}) + h_1^j(x_t,\, a_t,\, s_t) \geq 0, \;\; j = 0, ..., l, \;\; t \geq 1
\end{aligned}
$$

for initial conditions $x_0 = 0$ and $s_0 = s$, where φ is defined as

$$
\varphi^j(\mu,\, \gamma) \equiv \begin{array}{ll} \mu^j + \gamma^j & \text{if } \; N_j = \infty, \; \text{i.e. } \; j = 0, ..., k \\ \gamma^j & \text{if } \; N_j = 0, \; \text{i.e. } \; j = k+1, ..., l \end{array}
$$

The usefulness of $\mathbf{SPP}_{\mu}$ comes from the fact that its objective function has a very special form: the term inside the expectation in (**??**) is precisely the objective function of $\mathbf{PP}_{\varphi(\mu,\, \gamma^*)}$ given the states (x_1, s_1). This will allow us to show that if $(\{a_t^*\}_{t=0}^{\infty}, \gamma^*)$ solves $\mathbf{SPP}_{\mu}$ at (x_0, s_0) then $\{a_{t+1}^*\}_{t=0}^{\infty}$ solves $\mathbf{PP}_{\varphi(\mu,\, \gamma^*)}$ at (x_1, s_1), where $x_1 = \ell(x_0,\, a_0^*,\, s_1)$; that is , the continuation problem that needs to be solved in the next period is precisely a planner problem where the weights have been shifted according to φ.

We then show that, under fairly general conditions, solutions to $\mathbf{PP}_{\mu}$ are solutions to $\mathbf{SPP}_{\mu}$ (Theorem 1), and viceversa (Theorem 2). Also, the usual slackness conditions will guarantee that if $(\{a_t^*\}, \gamma^*)$ solves $\mathbf{SPP}_{\mu}$

$$
\mathrm{E}_0 \sum_{j=0}^{l} \gamma^{j*} \left[\sum_{t=1}^{N_j+1} \beta^t h_0^j(x_t^*,\, a_t^*,\, s_t) + h_1^j(x,\, a_0^*,\, s)\right] = 0, \tag{9}
$$

so that the values of achieved by $\mathbf{SPP}_{\mu}$ and $\mathbf{PP}_{\mu}$ at their solutions will coincide.

If $\mathbf{PP}_{\mu}$ was a standard dynamic programming problem, then the following Bellman equation would be satisfied

$$
\begin{aligned}
V_{\mu}(x, s) \quad = \quad & \sup \{\mu h_0(x,\, a,\, s) + \beta \mathrm{E}_0 V_{\mu}(x^{*\prime}, s')\} \\
& \text{s.t .(\textbf{??})}
\end{aligned} \tag{10}
$$

The reason is that in standard dynamic programming if the problem is reoptimized at period $t = 1$ given the states (x_1, s_1), the reoptimization simply confirms the choice that had been previously made for $t \geq 1$. However, with forward looking constraints (**??**) as in $\mathbf{PP}_{\mu}$ this Bellman equation is *not* satisfied, the reason being that if the problem is reoptimized at $t = 1$ the choice

[6]We use the notation $\mu h_0(x, a, s) \equiv \sum_{j=0}^{l} \mu^j h_0^j(x, a, s)$.

will violate the forward looking constraint of period $t = 0$, if (**??**) is binding at $t = 0$. A central element of our approach is that, as suggested by the objective function of $\mathbf{SPP}_\mu$, if the solution is reoptimized in period $t = 1$ with the new weights $\varphi(\mu, \gamma^*)$ we still have a solution of the original problem $\mathbf{PP}_\mu$. This allows the construction of a recursive formulation to our original $\mathbf{PP}_\mu$ problem where the value function with modified weights is included in the right side of the functional equation, to capture the terms in (**??**). More specifically, we show that under fairly general assumptions solutions to $\mathbf{SPP}_\mu$ obey a saddle-point functional equation **SPFE**

$$\textbf{SPFE} \qquad W(x,\, \mu,\, s) = \inf_{\gamma \geq 0} \sup_{a \in A} \left\{ \mu h_0(x,\, a,\, s) + \gamma h_1(x,\, a,\, s) + \beta \, \mathrm{E}\left[W(x',\, \mu',\, s') \middle| \, s \right] \right\}$$
$$\text{s.t. } x' = \ell(x,\, a,\, s'),\ p(x, a,\, s) \geq 0$$
$$\text{and } \mu' = \varphi(\mu,\, \gamma).$$

when the value function is $W(x,\, \mu,\, s) = V_\mu(x, s)$

To simplify the exposition we will only consider problems where policy choices are uniquely defined and, therefore, associated with a value function W satisfying **SPFE** there is *a policy function*}ψ; i.e. $(\, a^*,\, \gamma^*) = \psi((x, \mu, s)$.

Once we show that $V_\mu(x, s)$ satisfies the **SPFE** (Theorem 3), we show that if $W(x, \mu, s)$ satisfies **SPFE,** and its solutions are unique, the corresponding path, $\{a_t^*, \gamma_t^*\}$ generated by ψ, is a solution to $\mathbf{SPP}_\mu$ in state (x, s); that is, $(\{a_t^*\}_{t=0}^\infty, \gamma_0^*)$ solves $\mathbf{SPP}_\mu$ in state (x, s), $(\{a_t^*\}_{t=1}^\infty, \gamma_1^*)$ solves $\mathbf{SPP}_{\mu^{*\prime}}$ in state $(x^{*\prime}, s')$, etc. (Theorem 4). More precisely, once we show that the value of $\mathbf{PP}_\mu$ at $(x, s), V_\mu(x, s)$, satisfies **SPFE,** we can extend the dynamic programming principle to show that the modified problem $\mathbf{PP}_\mu$ is the 'correct continuation problem' to the planners' problem, in the sense that, if μ is properly updated, the solution can be found each period by re-optimizing a version of $\mathbf{PP}_\mu$.

An implication of this result is that $\mathbf{PP}_\mu$ can be seen as a "continuation problem" that insures full commitment to the solution chosen at $t = 0$, in the sense that $\mathbf{PP}_{\varphi(\mu,\, \gamma^*)}$ is the problem that 'the committed planner should solve in order to follow the preannounced solution' if she were given the chance to reoptimize. The time-inconsistency problem arises from the fact that, if intertemporal incentive constraints of the form (**??**) are binding in period $t = 0$, a sequential decision reoptimizing $\mathbf{PP}_\mu$ in period 1, state (x_1^*, s_1), does *not* result in the optimal choice $\left\{a_{t+1}^*\right\}_{t=0}^\infty$. However, we show that $\left\{a_{t+1}^*\right\}_{t=0}^\infty$ is the solution to a properly modified planner's problem, namely, $\mathbf{PP}_{\varphi(\mu,\, \gamma^*)}$ in state (x_1^*, s_1), where γ^* is the Lagrange multiplier of $\mathbf{SPP}_\mu$. Therefore, if a planner re-optimizes using $\mathbf{PP}_{\varphi(\mu,\, \gamma^*)}$ in state (x_1^*, s_1), she chooses to follow the original plan $\left\{a_{t+1}^*\right\}_{t=0}^\infty$.

The main results of this paper can be summarized as saying that, under relatively standard conditions: *i)* a $\mathbf{PP}_\mu$ problem has a recursive **SPFE** representation, and *ii)* solutions to $\mathbf{PP}_\mu$ can be obtained by solving recursive saddle point problems **SPFE**.

Before we turn to these results in Sections 4 and 5, in the next Section we show how our approach is impelemented in a couple of canonical examples.

3 Some applications

In this Section, we illustrate our approach with two examples. In the first, there are only *intertemporal participation constraints* (i.e. $k = l$ in (**??**)); in the second, there are only *intertemporal one-period (Euler) constraints* (i.e. $k = 0$ in (**??**)). The first is similar to the model studied in Marcet and Marimon (1992), Kocherlakota (1996), Kehoe and Perri (2002), among others, and it is canonical of models with intertemporal default constraints; the second is based on the model studied by Aiyagari, Marcet, Sargent and Seppala (2002) and it is canonical models with Euler constraints, as in Ramsey models.

3.1 Intertemporal participation constraints.

We consider, as an example, a model of a partnership, where several agents can share their individual risks and jointly invest in a project which can not be undertaken by single (or subgroups of) agents. Formally, there is a single good and J infinitely-lived consumers, with preferences represented by $E_0 \sum_{t=0}^{\infty} \beta^t \ u(c_{j,t})$; u is assumed to be bounded, strictly concave and monotone, with $u(0) = 0$; c represents individual consumption. Agent j receives an endowment of consumption good $y_{j,t}$ at time t and, given a realization y_t, $y_t = \sum_{j=1}^{J} y_{j,t} > 0$, agent j has an outside option $v_j^a(y_t)$. For simplicity we consider that the outside option is the autarkic solution: $v_j^a(y_t) = \mathrm{E}\left[\sum_{n=0}^{\infty} \beta^n \ u(y_{j,t+n}) \mid y_{j,t}\right]$, which implicitly assumes that agent j is permanently excluded from the partnership and, if she defaults, does not have any claims on the capital of the partnership[7]Total production is given by $F(k, \theta)$, and it can be split into consumption c and investment i. The stock of capital k depreciates at the rate δ. The joint process $\{\theta_t, y_t\}_{t=0}^{\infty}$ is assumed to be Markovian and the initial conditions (k_0, θ_0, y_0) are given. The planner's problem takes the form:

$$\mathbf{PP} \qquad \max_{\{c_t, i_t\}} \ \mathrm{E}_0 \sum_{t=0}^{\infty} \beta^t \ \sum_{j=1}^{J} \alpha_j \ u(c_{j,t})$$

$$\text{s.t. } k_{t+1} = (1-\delta)k_t + i_t, \quad F(k_t, \theta_t) + y_t - \left(\sum_{j=1}^{J} c_{j,t} + i_t \right) \geq 0$$

$$\mathrm{E}_t \sum_{n=0}^{\infty} \beta^n \ u(c_{j,t+n}) \geq v_j^a(y_t) \qquad \text{for all } j, t \geq 0.$$

It is easy to map this planner's problem into our $\mathbf{PP}_\mu$ formulation. Let $s \equiv (\theta, y)$; $x \equiv k$; $a \equiv (i, c)$; $\ell(x,\, a,\, s) \equiv (1-\delta)k + i$; $p(x,\, a,\, s) \equiv F(k, \theta) + \sum_{j \in J} y_j -$

[7]For a model where there is no permanent exclusion and, therefore, the outside option can not be taken as exogenous, see Cooley, Marimon and Quadrini (2000). Their model requires to map outside options, taken as exogenous functions, into value functions of the corresponding recursive contracts, and then find the appropriate fixed point of that map.

$\left(\sum_{j\in J} c_j + i\right)$; $h_0^0(x,\, a,\, s) \equiv \sum_{j=1}^J \alpha_j\ u(c_j)$, $h_0^j(x,\, a,\, s) \equiv u(c_j), h_1^j(x,\, a,\, s) \equiv h_0^j(x,\, a,\, s) - v_j^a(y_t)$, $j = 1, ..., J$, and $h_1^0(x,\, a,\, s) \equiv h_0^0(x,\, a,\, s) - R$, where R is large enough as to guarantee $\sum_{j=1}^J \alpha_j\ u(c_j) - R > 0$. Finally, apply the convention $\mu_0 = (1, 0, ..., 0)$ and **PP** is just $\mathbf{PP}_{\mu_0}$. In its original notation, **SPFE** takes the form

$$
\begin{aligned}
W(k,\, \mu, y, \theta) &= \inf_{\gamma\geq 0} \sup_{c,i} \{\sum_{j=1}^J \left((\mu_0 \alpha_j + \mu_j)\ u(c_j) - \gamma^j \left(u(c_j) - v_j^a(y) \right) \right) \\
&+ \gamma^0 \left(\sum_J \alpha_j u_j(c_j) - R \right) + \beta\ \mathrm{E}\left[W(k',\, \mu', y', \theta') \,|\omega, \theta \right] \} \\
&\text{s.t. } k' = (1-\delta)k + i, \quad F(k,\theta) + \sum_{j=1}^J y_j - \left(\sum_{j=1}^J c_j + i \right) \geq 0 \\
&\text{and } \mu' = \mu + \gamma
\end{aligned}
$$

Letting ψ be the policy function associated with this functional equation, efficient allocations satisfy $(c_t,\, i_t, \gamma_t) = \psi(k_t,\, \mu_t,\, \theta_t,\, y_t)$ with initial conditions $(k_0, \overline{\mu}_0,\, \theta_0,\, y_0)$.

Notice that solutions to **SPP** satisfy $\mu_t^0 = 1$. It follows that co-state variables μ become the weights that the planner assigns to each agent, which evolve according to whether or not the participation constraint is binding. Furthermore, if u is differentiable,

$$
\frac{u'(c_{i,t})}{u'(c_{j,t})} = \frac{\alpha_j + \mu_{t+1}^j}{\alpha_i + \mu_{t+1}^i} \quad \text{, for all } i, j \text{ and } t.
$$

Thus, the optimal allocations amount to choosing efficiently the time profile of the time-dependent weights $(\alpha_j + \mu_{j,t+1})$, in such a way that the participation constraints are satisfied. Every time that the participation constraint for an agent is binding, his weight is increased by the amount of the corresponding lagrange multiplier. An agent is induced not to default by increasing his consumption not only in the period where he is tempted to default, but also for many of the following periods; in this way, the additional consumption that the agent receives to prevent default is smoothed over time. That is, individual paths of consumption depend on individual histories (in particular, on past "temptations to default") not just on the initial wealth distribution and the aggregate consumption path, as in the Arrow-Debreu competitive allocations. This also shows that if enforcement constraints are never binding (e.g., punishments are severe enough) then $\mu_t = \mu_0$ and we recover the "constancy of the marginal utility of expenditure", and the "constant proportionality between individual consumptions," given by $u'(c_{i,t})/u'(c_{j,t}) = \alpha_j/\,\alpha_i$. In other words, the evolution of the co-state variables can be also interpreted as the evolution of the distribution of wealth.

The intertemporal Euler equation of $\mathbf{SPP}_\mu$ is also very informative:

$$\mu^i_{t+1} u'(c_{i,t}) = \beta \mathrm{E}\left[\mu^i_{t+2} u'(c_{i,t+1}) \left(F_{k_{t+1}} + (1-\delta)\right)\right]$$

where, as with $j = 0, ..., k$, constraints: $\mu^i_{t+2} = \mu^i_{t+1} + \gamma^i_{t+1}$. That is, the 'stochastic discount factor', $\beta u'(c_{i,t+1})/u'(c_{i,t})$ is distorted by $\left(1 + \gamma^i_{t+1}/\mu^i_{t+1}\right)$, a distortion which does not vanish unless the non-negative process $\left\{\gamma^i_t\right\}$ converges to zero. But if intetemporal participation constraints are infinitely often binding, there will be a non degenerate distribution of consumption in the long-run; in contrast with an economy where intetemporal participation constraints cease to be binding, as in an economy with full enforcement.

3.2 Intertemporal implementability (Euler equation) constraints

Example 2. A Ramsey problem

We briefly sketch a version of the optimal taxation problem studied by Aiyagari, Marcet, Sargent and Seppala (2002) [to be explained better]. A representative consumer solves

$$\begin{aligned} & \max E_0 \sum_{t=0}^{\infty} \delta^t \quad [u(c_t) + v(l_t)] \\ \text{s.t.} \qquad & c_t + b_{t+1} p^b_t = l_t(1-\tau_t) + b_t \end{aligned}$$

The government must finance exogenous random expenditures g issuing debt and collecting taxes. Given the technology $c_t + g_t = l_t$,the budget of the government mirrors the budget of the representative agent and is subject to a no-Ponzi constraint. In a competitive equilibrium, the following intertemporal and intratemporal equations must be satisfied:

$$\begin{aligned} p^b_t u'(c_t) &= \beta \; E_t u'(c_{t+1}) \\ -\frac{v'(l_t)}{u'(c_t)} &= (1-\tau_t). \end{aligned}$$

Therefore, the Ramsey problem can be formulated as

$$\begin{aligned} & \max_{\{c_t, b_{t+1}\}} E_0 \sum_{t=0}^{\infty} \beta^t \; [u(c_t) + v(l_t)] \\ \text{s.t.} \qquad & c_t \; + b_{t+1} \beta \; E_t \frac{u'(c_{t+1})}{u'(c_t)} = -l_t \frac{v'(l_t)}{u'(c_t)} + b_t \end{aligned}$$

This problem can be represented as a $\mathbf{PP}_{\mu_0}$ by letting: $s \equiv g$; $x \equiv b$; $a \equiv (c, y)$; $p(x, a, s) \equiv l - (c + g)$, $h^0_0(x, a, s) \equiv u(c_t) + v(l_t)$, $h^1_0(x, a, s) \equiv u'(c_{t+1})$, $h^0_1(x, a, s) \equiv h^0_0(x, a, s) - R$ (big); $h^1_1(x, a, s) \equiv y$, where

$$y_t = \frac{-l_t v'(l_t) + b_t u'(c_t) - c_t u'(c_t)}{b_{t+1}} \; .$$

In particular, $\mathbf{SPP}_\mu$ solutions satisfy

$$\mathrm{E}\left[(\mu^1_{t+1} - \gamma^1_{t+1})u'(c_{t+1})\right] = 0 \tag{11}$$

where, as with $j = k+1, ..., l$, constraints, $\mu^1_{t+1} = \gamma^1_t$. As it can be seen, (**??**) immediately shows the nature of the distortion – that is, of the 'time-inconsistency' problem – and of its possible resolution: the convergence of the random variable γ^1_t. In particular, Aiyagari et al. show that, in fact, the process $\left\{\gamma^1_t\right\}$ is a non-negative submartingale.

4 The relationship between $\mathbf{PP}_\mu$, $\mathbf{SPP}_\mu$, and SPFE

This section proves the relationships between the initial maximization problem $\mathbf{PP}_\mu$, the saddle-point problem $\mathbf{SPP}_\mu$ and the saddle-point functional equation **SPFE** discussed in the previous Sections. We first describe the basic structure of the problems being considered.

4.1 Basic Structure

There exist an exogenous stochastic process $\{s_t\}_{t=0}^{\infty}$, $s_t \in S$, defined on the probability space $(S_\infty, \mathcal{S}, P)$. As usual, s^t denotes a history $(s_0, ..., s_t)$ and $\mathcal{S}_t$ the sigma-algebra of events of s^t. An action in period t, history s^t, is denoted by $a_t(s^t)$, where $a_t(s^t) \in A \subset R^m$; when there is no confusion, it is simply denoted by a_t. Given s_t and the endogenous state $x_t \in X \subset R^n$, an action a_t is (technologically) feasible if $p(x_t, a_t, s_t) \geq 0$. If the latter feasibility condition is satisfied, the endogenous state evolves according to $x_{t+1} = \ell(x_t,\, a_t,\, s_{t+1})$. Plans, $\boldsymbol{a} = \{a_t\}_{t=0}^{\infty}$, are elements of $\mathcal{A} = \{\boldsymbol{a} : \forall t \geq 0,\, a_t : S_t \to A$ and $a_t \in \mathcal{L}^m_\infty(S_t, \mathcal{S}_t, P),\, \}$, where $\mathcal{L}^m_\infty(S_t, \mathcal{S}_t, P)$ denotes the space of m-valued, essentially bounded, $\mathcal{S}_t$-measurable functions. The corresponding endogenous state variables are elements of $\mathcal{X} = \{\boldsymbol{x} : \forall t \geq 0,\, x_t \in \mathcal{L}^n_\infty(S_t, \mathcal{S}_t, P)\}$.

Given initial conditions (x, s), a plan $\boldsymbol{a} \in \mathcal{A}$ and the corresponding $\boldsymbol{x} \in \mathcal{X}$, the evaluation of such plan in $\mathbf{PP}_\mu$ is given by

$$f_{(x,\mu.s)}(\boldsymbol{a}) = \mathrm{E}_0 \sum_{j=0}^{k} \sum_{t=0}^{N_j} \beta^t \mu^j h^j_0(x_t,\, a_t,\, s_t)$$

We can describe the forward-looking constraints by defining $g : \mathcal{A} \to \mathcal{L}^{k+1}_\infty$ coordinatewise as

$$g(\boldsymbol{a})^{\,j}_{\,t} = \mathrm{E}_t \left[\sum_{n=1}^{N_j+1} \beta^n h^j_0(x_{t+n},\, a_{t+n},\, s_{t+n}) \right] + h^j_1(x_t,\, a_t,\, s_t)$$

Given initial conditions (x, s), the corresponding feasible set of plans is then

$$\begin{aligned} \mathcal{B}(x,s) \;\; = \;\; & \{\boldsymbol{a} \in \mathcal{A} :\; p(x_t, a_t,\, s_t) \geq 0,\; g(\boldsymbol{a})_{\,t} \geq 0,\, \boldsymbol{x} \in \mathcal{X}\;, \\ & x_{t+1} = \ell(x_t,\, a_t,\, s_{t+1}) \text{ for all } t \geq 0,\; \text{given } (x_0, s_0) = (x, s)\}\,. \end{aligned}$$

Then $\mathbf{PP}_\mu$ can be written in compact form as

$$\mathbf{PP}_\mu \qquad \sup_{\boldsymbol{a}\in\mathcal{B}(x,s)} f_{(x,\mu.s)}(\boldsymbol{a})$$

We denote solutions to this problem as $\boldsymbol{a}^*$ and the corresponding sequence of state variables $\boldsymbol{x}^*$. When the solution exists we define the value function of $\mathbf{PP}_\mu$ as

$$V_\mu(x,s) = f_{(x,\mu.s)}(\boldsymbol{a}^*) \tag{12}$$

Similarly, we can also write $\mathbf{SPP}_\mu$ in a compact form, by defining

$$\begin{aligned}\mathcal{B}'(x,s) \;\; &= \;\; \{\boldsymbol{a}\in\mathcal{A} : p(x_t, a_t,\, s_t) \geq 0,\, g(\boldsymbol{a})_{\,t+1} \,\geq 0;\; \boldsymbol{x}\in\mathcal{X} \\ &\qquad x_{t+1} = \ell(x_t,\, a_t,\, s_{t+1}) \text{ for all } t\geq 0,\; \text{given } (x_0,s_0)=(x,s)\}\,.\end{aligned}$$

$$\mathbf{SPP}_\mu \qquad \inf_{\gamma\in R^l_+}\sup_{\boldsymbol{a}\in\mathcal{B}'(x,s)} \left\{f_{(x,\mu,.s)}(\boldsymbol{a}) + \gamma g(\boldsymbol{a})_0\right\}$$

Note that $\mathcal{B}'$ only differs from $\mathcal{B}$ in that the forward-looking constraints in period zero $g(\boldsymbol{a})_{\,0}\,\geq 0$ are not included as a condition in the set $\mathcal{B}'$, instead these constraints form part of the objective function of $\mathbf{SPP}_\mu$.

4.2 Assumptions and existence of solutions to $\mathbf{PP}_\mu$

We consider the following set of assumptions:

A1. s_t takes values on a set $S\subset R^K$. $\{s_t\}_{t=0}^\infty$ is a Markovian stochastic process defined on the probability space $(S_\infty,\mathcal{S},P)$.

A2. **(a)** $X\subset R^n$ and A is a closed subset of R^m. **(b)** The functions $p : X\times A\times S\to R$ and $\ell : X\times A\times S\to X$ are measurable and continuous.

A3. Given (x,s),there exist constants $B>0$ and $\varphi\in(0,\beta^{-1})$, such that if $p(x,a,s)\geq 0$ and $x'=\ell(x,\,a,s')$, then $\|a\|\leq B\,\|x\|$ and $\|x'\|\leq\varphi\,\|x\|$

A4. The functions $h_i^j(\cdot,\,\cdot,\,s)$, $i=0,1$, $j=0,...,l$, are continuous and uniformly bounded, and $\beta\in(0,1)$.

A5. The function $\ell(\cdot,\,\cdot,s)$ is linear and the function $p(\cdot,\,\cdot,s)$ is concave. X and A are convex sets.

A6. The functions $h_i^j(\cdot,\,\cdot,\,s)$, $i=0,1$, $j=0,...,l$, are concave.

A6s. In addition to **A6**, the functions $h_0^j(x,\,\cdot,\,s), j=0,...,l$, are strictly concave.

A7. For all $(x,\,s)$, there exists a program $\{\widetilde{a}_n\}_{n=0}^\infty$, with initial conditions (x,s), which satisfies the inequality constraints (**??**) and (**??**) with strict inequality.

Assumptions **A1** and **A2** are part of our basic structure, described in the previous sub-section. These asumptions, together with **A3-A4,** are standard and we treat them as our basic assumptions. Assumptions **A5-A7** are often made but they are not satisfied in some interesting models; however, these assumptions are only used in some of the results below. For example, the concavity assumptions **A5-A6** are not needed for many results, and assumption **A7** is a standard interiority assumption, only needed to guarantee the existence of Lagrange multipliers.

The following proposition gives sufficient conditions for a maximum to exist for any μ. The aim is not to have the most general existence theorem[8], but to stress that one can find fairly general conditions under which $\mathbf{PP}_\mu$ has a solution for any μ, which will be crucial in the discussion of how our approach compares with that of Abreu Pearce and Stachetti, since this ensures that the continuation problem (namely $\mathbf{PP}_{\varphi(\mu,\gamma)}$) is well defined for any γ.

Proposition 1. Assume **A1-A6** and that the set of possible exogenous states S is countable. Fix $(x, \mu, s) \in X \times R_+^{l+1} \times S$. Assume there exists a feasible plan $\widetilde{\boldsymbol{a}} \in \mathcal{B}(x, s)$ such that $f_{(\mu,x,s)}(\widetilde{\boldsymbol{a}}) > -\infty$. Then there exists a program $\boldsymbol{a}^*$ which solves $\mathbf{PP}_\mu$ with initial conditions $x_0 = x$, $s_0 = s$.

Furthermore, if **A6s** is also satisfied then the solution is (almost sure) unique.

Proof: See Appendix.

4.3 The relationship between $\mathbf{PP}_\mu$ and $\mathbf{SPP}_\mu$

The following result says that a solution to the maximum problem is also a solution to the saddle point problem. It follows from standard theory of constrained optimization in linear vector spaces (see, for example, Luenberger (1969, Section 8.3, Theorem 1 and Corollary 1). As in the standard theory, convexity and concavity assumptions (**A5 to A6**), as well as an interiority assumption (**A7**) are necessary in order to obtain the result.

Theorem 1 ($\mathbf{PP}_\mu \implies \mathbf{SPP}_\mu$). Assume **A1-A7** and fix $\mu \in R_+^{l+1}$. Let $\boldsymbol{a}^*$ be a solution to $\mathbf{PP}_\mu$ with initial conditions (x, s).There exists a $\gamma^* \in R_+^l$ such that $(\boldsymbol{a}^*, \gamma^*)$ is a solution to $\mathbf{SPP}_\mu$ with initial conditions (x, s).

Furthermore, the value of $\mathbf{SPP}_\mu$ is the same as the value of $\mathbf{PP}_\mu$, more precisely

$$V_\mu(x, s) = f_{(x,\mu,.s)}(\boldsymbol{a}^*) + \gamma^* g(\boldsymbol{a}^*)_0 \tag{13}$$

Proof: It is an immediate application of Theorem 1 (8.3) in Luenberger (1969), p. 217.

The following is a theorem on the sufficiency of a saddle point for a maximum.

[8] For example, not requiring **A6** or the countability of S.

Theorem 2 ($\mathbf{SPP}_\mu \Longrightarrow \mathbf{PP}_\mu$). Given any $(x, s) \in X \times R_+^{l+1} \times S$, let $(\boldsymbol{a}^*, \gamma^*)$ be a solution to $\mathbf{SPP}_\mu$ for initial conditions (x, s). Then $\boldsymbol{a}^*$ is a solution to $\mathbf{PP}_\mu$ for initial conditions $(x,\, s)$.

Furthermore, the value of the two programs is the same and (**??**) holds.

Notice that Theorem 2 is a sufficiency theorem 'almost free of assumptions.' All that is needed is the basic structure of section **??** defining the corresponding infinite-dimensional optimization and saddle-point problems *together with* the assumption that a solution to $\mathbf{SPP}_\mu$ exists. Once these conditions are satisfied assumptions **A2** to **A7** are not needed.

Proof: The following proof is an adaptation, to $\mathbf{SPP}_\mu$, of a sufficiency theorem for Lagrangian saddle points (see, for example, Luenberger (1969), Theorem 8.4.2, p.221).

If $(\boldsymbol{a}^*,\, \gamma^*)$ solves $\mathbf{SPP}_\mu$, minimality of γ^* implies that, for every $\gamma \geq 0$,

$$(\,\gamma^* + \gamma)\, g(\boldsymbol{a}^*)_0 \geq \gamma^* g(\boldsymbol{a}^*)_0;$$

therefore, $g(\boldsymbol{a}^*)_0 \geq 0$, but since $\boldsymbol{a}^* \in \mathcal{B}'(x, s)$, it follows that $\boldsymbol{a}^* \in \mathcal{B}(x, s)$; i.e. $\boldsymbol{a}^*$ is a feasible program for $\mathbf{PP}_\mu$. Furthermore, the minimality of γ^* implies that

$$\gamma^* g(\boldsymbol{a}^*)_0 \leq 0 g(\boldsymbol{a}^*)_0 = 0$$

but since $\gamma^* \geq 0$ and $g(\boldsymbol{a}^*)_0 \geq 0$, it follows that $\gamma^* g(\boldsymbol{a}^*)_0 = 0$. Now, suppose there exist $\tilde{\boldsymbol{a}} \in \mathcal{B}(x, s)$ satisfying $f_{(x,\mu,s)}(\tilde{\boldsymbol{a}}) > f_{(x,\mu,s)}(\boldsymbol{a}^*)$, then, since $\gamma^* g(\tilde{\boldsymbol{a}})_0 \geq 0$, it must be that

$$f_{(x,\mu,s)}(\tilde{\boldsymbol{a}}) + \gamma^* g(\tilde{\boldsymbol{a}})_0 > f_{(x,\mu,s)}(\boldsymbol{a}^*) + \gamma^* g(\boldsymbol{a}^*)_0$$

which contradicts the maximality of $\boldsymbol{a}^*$ for $\mathbf{SPP}_\mu$.

Finally, using $\gamma^* g(\boldsymbol{a}^*)_0 = 0$, we have $f_{(x,\mu,s)}(\boldsymbol{a}^*) + \gamma^* g(\boldsymbol{a}^*)_0 = V_\mu(x, s)$ ■

4.4 The relationship between $\mathbf{SPP}_\mu$, and SPFE

Recall that a function $W : X \times R_+^{l+1} \times S \to R$ satisfies **SPFE** at $(x,\, \mu,\, s)$ when

$$W(x,\ \mu,\ s) = \min_{\gamma \geq 0} \max_{a \in \mathcal{A}} \{\mu h_0(x,\, a,\, s) + \gamma h_1(x,\, a,\, s) + \beta\, \mathrm{E}\,[W(x',\, \mu',\, s')|\, s]\} \tag{14}$$

$$\text{s.t. } x' = \ell(x,\, a,\, s), \quad p(x, a,\, s) \geq 0 \tag{15}$$

$$\text{and } \mu' = \varphi(\mu,\, \gamma). \tag{16}$$

In (**??**) we have substituted inf sup by min max, implicitely assuming that a solution to the saddle point problem exists, in which case the value, $W(x,\, \mu,\, s)$, is uniquely determined[9]. In other words, the right side of **SPFE** is well defined for all (x, μ, s) and W for which a saddle point exists.

We say that W satisfies **SPFE** if it satisfies **SPFE** at *any possible state* $(x, \mu, s) \in X \times R_{+}^{l+1} \times S$. Given W,we define the *saddle-point policy correspondence* (*SP policy correspondence*) $\Psi : X \times R_{+}^{l+1} \times S \to A \times R_{+}^{l+1}$ by

$$\begin{aligned}\Psi_W(x,\, \mu,\, s) \quad = \quad & \Big\{ (\, a^*, \gamma^*) : a^* \in \arg\max_{a\in\mathcal{A},\; x'\in X} \mu h_0(x,\, a,\, s) + \gamma^* h_1(x,\, a,\, s) + \beta \,\mathrm{E}\left[W(x',\, \mu^{*\prime},\, s') |\, s\right] \\ & \text{for } \mu^{*\prime} = \varphi(\mu,\, \gamma^*) \text{ and } \;(\textbf{??}); \\ & \gamma^* \in \arg\min_{\gamma\geq 0} \mu h_0(x,\, a^*,\, s) + \gamma h_1(x,\, a^*,\, s) + \beta \,\mathrm{E}\left[W(x^{*\prime},\, \mu',\, s') |\, s\right], \\ & \text{for } x^{*\prime} = \ell(x,\, a^*,\, s) \text{ and } (\textbf{??})\Big\}\end{aligned}$$

If Ψ_W is single valued, we denote it by ψ_W, and we call it a *saddle-point policy function* (*SP policy function*).

We define the function $W^*(x, \mu, s) \equiv V_\mu(x, s)$. The following theorem says that W^* satisfies **SPFE.**

Theorem 3 (SPP$_\mu \Longrightarrow$ SPFE). Assume that **SPP**$_\mu$ has a solution for any $(x, \mu, s) \in X \times R_{+}^{l+1} \times S$, then W^* satisfies **SPFE**. Furthermore, letting $(\boldsymbol{a}^*,\, \gamma^*)$ be a solution to **SPP**$_\mu$ at (x, s), we have $(a_0^*, \gamma^*) \in \Psi_{W^*}(x,\, \mu,\, s)$.

As in Theorem 2, Theorem 3 is also a theorem 'almost free of assumptions,' once the underlying structure and the existence of a well defined solution to **SPP**$_\mu$ at all possible (x, μ, s) is assumed.

Proof: By theorem 2, we have that whenever **SPP**$_\mu$ has a solution W^* is well defined. Then, we first prove that, for any given (x, μ, s), if $(\boldsymbol{a}^*,\, \gamma^*)$ solves **SPP**$_\mu$ at (x, s) the following recursive equation is satisfied:

$$W^*(x, \mu, s) = \mu h_0(x,\, a_0^*,\, s) + \gamma^* h_1(x,\, a_0^*,\, s) + \beta \,\mathrm{E}\left[W^*(x_1^*,\, \varphi(\mu,\, \gamma^*),\, s') |\, s\right] \tag{17}$$

To prove $\leq$ in (**??**) we write

$$\begin{aligned} W^*(x, \mu, s) \quad &= \quad f_{(x,\mu,s)}(\boldsymbol{a}^*) + \gamma^* g(\boldsymbol{a}^*)_0 \\ &= \quad \mu h_0(x,\, a_0^*,\, s) + \gamma^* h_1(x,\, a_0^*,\, s) \\ & \qquad + \beta \mathrm{E}\left[\sum_{j=0}^{l} \varphi^j(\mu,\, \gamma^*) \sum_{t=0}^{N_j} \beta^t\; h_0^j(x_{t+1}^*,\, a_{t+1}^*,\, s_{t+1}) \mid s \right] \\ &= \quad \mu h_0(x,\, a_0^*,\, s) + \gamma^* h_1(x,\, a_0^*,\, s) + \beta \mathrm{E}\left[f_{(x_1^*, \varphi(\mu,\, \gamma^*), s_1)}(\sigma \boldsymbol{a}^*) |\, s\right] \\ &\leq \quad \mu h_0(x,\, a_0^*,\, s) + \gamma^* h_1(x,\, a_0^*,\, s) + \beta \mathrm{E}\left[V_{\varphi(\mu,\, \gamma^*)}(x_1^*, s_1) |\, s\right] \\ &= \quad \mu h_0(x,\, a_0^*,\, s) + \gamma^* h_1(x,\, a_0^*,\, s) + \beta \mathrm{E}\left[W^*(x_1^*, \varphi\,(\mu,\, \gamma^*, s)) |\, s\right] \end{aligned}$$

where $\sigma\boldsymbol{a}^*$ is the original optimal sequence shifted one period; formally, letting the shift operator $\sigma : S^{t+1} \to S^t$ be given by $\sigma(s^t) = (s_1, s_2 ..., s_t)$,

[9]See Lemma 3A in Appendix B.

we define the $\mathcal{S}_{t+1}$-measurable function σa_t^* by $\sigma a_t^*(s) \equiv a_{t+1}^*(s)$. The first equality follows from the definition of W^* and because Theorem 2 guarantees (**??**), the second equality follows from the definition of f, g and simple algebra, the third equality follows from the definitions of f, φ, and $\boldsymbol{a}^*$. The weak inequality follows from the fact $\boldsymbol{a}^*$ is a feasible solution for the problem $\mathbf{PP}_{\varphi(\mu,\,\gamma^*)}$ with initial conditions (x_1^*, s_1) and that this program achieves $V_{\varphi(\mu,\,\gamma^*)}(x_1^*, s_1)$ at its maximum, and the last equality follows from Theorem 2 and (**??**).

To show $\geq$ (**??**) we construct a sequence $\boldsymbol{a}^+$ that consists of the optimal choice for $\mathbf{SPP}_\mu$ for initial conditions (x, s) in the initial period, but subsequently is followed by the optimal choices for $\mathbf{PP}_{\varphi(\mu,\,\gamma^*(x,\mu,s))}$ for initial conditions (x_1^*, s_1). To define $\boldsymbol{a}^+$ formally we explicitly denote by $(\boldsymbol{a}^*(x,\mu,s), \gamma^*(x,\mu,s))$ a solution to $\mathbf{SPP}_\mu$ for given initial conditions (x, s) and we let

$$\begin{aligned}
a_0^+(x,\mu,s) &= a_0^*(x,\mu,s) \\
a_t^+(x,\mu,s) &= \sigma a_{t-1}^*(x_1^*, \varphi\left(\mu,\, \gamma^*(x,\mu,s), s\right),\ s_1)
\end{aligned}$$

for all (x,μ,s) and $t \geq 1$. Also, we let $\boldsymbol{x}^+$ be the corresponding sequence of state variables.

In what follows, we simplify again notation and we go back to denoting $a_t^*(x,\mu,s)$ by a_t^*, $\gamma^*(x,\mu,s)$ by γ^*, and $a_t^+(x,\mu,s)(s^t)$ by a_t^+; then, we have:

$$\begin{aligned}
W^*(x,\mu,s) &= f_{(x,\mu,s)}(\boldsymbol{a}^*) + \gamma^* g(\boldsymbol{a}^*)_0 \\
&\geq f_{(x,\mu,s)}(\boldsymbol{a}^+) + \gamma^* g(\boldsymbol{a}^+)_0 \\
&= \mu h_0(x,\, a_0^+,\, s) + \gamma^* h_1(x,\, a_0^+,\, s) \\
&\quad + \beta \mathrm{E}\left[\sum_{j=0}^{l} \varphi(\mu,\, \gamma^*)^j \sum_{t=0}^{N_j} \beta^t\ h_0^j(x_{t+1}^+,\, a_{t+1}^+,\, s_{t+1}) \mid s\right] \\
&= \mu h_0(x,\, a_0^*,\, s) + \gamma^* h_1(x,\, a_0^*,\, s) \\
&\quad + \beta \mathrm{E}\left[W^*(x_1^*, \varphi\left(\mu,\, \gamma^*\right), s_1) | \, s\right],
\end{aligned}$$

where the first equality has been argued before, the first inequality follows from the fact that $\boldsymbol{a}^+(x,\mu,s)$ is a feasible allocation in $\mathbf{SPP}_\mu$ for initial conditions (x, s) – but $\boldsymbol{a}^*(x,\mu,s)$ is a solution of the max part to $\mathbf{SPP}_\mu$ at (x, s). The second equality just applies the definition of f and g, the last equality follows because $\boldsymbol{a}^+$ is optimal for $\mathbf{PP}_{\varphi(\mu,\,\gamma^*(x,\mu),s)}$ given initial conditions (x_1^*, s_1) from period 1 onwards and because Theorem 2 insures that (**??**) holds.

Notice that for this step of the proof it is crucial that we use $\mathbf{SPP}_\mu$ in order to obtain a recursive formulation. The first inequality above only works because we are considering a saddle point problem. Indeed, the $\boldsymbol{a}^+$ sequence (that reoptimizes in period $t = 1$) is feasible for $\mathbf{SPP}_\mu$ because this problem does not impose the forward looking constraints in $t = 0$. The sequence $\boldsymbol{a}^+$ would not be feasible in the original problem $\mathbf{PP}_\mu$, because by reoptimizing at period $t = 1$ the forward-looking constraints at $t = 0$ would be typically violated.

This ends the proof of (**??**).

To show that W^* satisfies **SPFE** we now prove that the right side of SPFE is well defined at W^* and that (a_0^*, γ^*) is a saddle point of the right side of (**??**) or, formally, that $(a_0^*, \gamma^*) \in \Psi_{W^*}(x, \mu, s)$.

We first prove that a_0^* solves the max part of the right side of SPFE. Given any $\widetilde{a} \in A, p(x, \widetilde{a}, s) \geq 0$, letting $\widetilde{a}_t^*(s^t) \equiv a_{t-1}^*(\ell(x, \widetilde{a}, s'), \varphi(\mu, \gamma^*), s')(\sigma(s^t))$ for $t \geq 1$, the definition of $\widetilde{a}_t^*$, and (**??**) give the first equality in

$$\begin{aligned}
& \mu h_0(x, \widetilde{a}, s) + \gamma^* h_1(x, \widetilde{a}, s) + \beta \mathrm{E}\left[W^*(\ell(x, \widetilde{a}, s'), \varphi(\mu, \gamma^*), s') | s\right] \\
= \; & \mu h_0(x, \widetilde{a}, s) + \gamma^* h_1(x, \widetilde{a}, s) \\
& + \beta \mathrm{E}\left[\sum_{j=0}^{l} \varphi(\mu, \gamma^*)^j \sum_{t=0}^{N_j} \beta^t \; h_0^j(\widetilde{x}_{t+1}^*, \widetilde{a}_{t+1}^*, s_{t+1}) \mid s\right] \\
= \; & f_{(x,\mu,s)}(\widetilde{\boldsymbol{a}}^*) + \gamma^* g(\widetilde{\boldsymbol{a}}^*)_0 \\
\leq \; & f_{(x,\mu,s)}(\boldsymbol{a}^*) + \gamma^* g(\boldsymbol{a}^*)_0 \\
= \; & W^*(x, \mu, s)
\end{aligned}$$

the second equality follows by definition, and the inequality because $(\boldsymbol{a}^*, \gamma^*)$ solves the max part of $\mathbf{SPP}_\mu$ while the third equality follows from (**??**). Now we can combine this with (**??**) to obtain that for all feasible $\widetilde{a} \in A$

$$\begin{aligned}
& \mu h_0(x, \widetilde{a}, s) + \gamma^* h_1(x, \widetilde{a}, s) + \beta \mathrm{E}\left[W^*(x_1^*, \varphi(\mu, \gamma^*), s') | s\right] \\
\leq \; & \mu h_0(x, a_0^*, s) + \gamma^* h_1(x, a_0^*, s) + \beta \mathrm{E}\left[W^*(x_1^*, \varphi(\mu, \gamma^*), s') | s\right]
\end{aligned}$$

implying that a_0^* solves the max part of the right side of the **SPFE**.

A similar argument shows that γ^* solves the min part**:** for any $\widetilde{\gamma} \in R_+^{l+1}$ let now

$$\begin{aligned}
& \mu h_0(x, a_0^*, s) + \widetilde{\gamma} h_1(x, a_0^*, s) + \beta \mathrm{E}\left[W^*(x_1^*, \varphi(\mu, \widetilde{\gamma}), s') | s\right] \\
\geq \; & \mu h_0(x, a_0^*, s) + \widetilde{\gamma} h_1(x, a_0^*, s) + \\
& + \beta \mathrm{E}\left[\sum_{j=0}^{l} \varphi(\mu, \widetilde{\gamma})^j \sum_{t=0}^{N_j} \beta^t \; h_0^j(x_{t+1}^*, a_{t+1}^*, s_{t+1}) \mid s\right] \\
= \; & f_{(x,\mu,s)}(\boldsymbol{a}^*) + \widetilde{\gamma} \; g(\boldsymbol{a}^*)_0 \\
\geq \; & f_{(x,\mu,s)}(\boldsymbol{a}^*) + \gamma^* \; g(\boldsymbol{a}^*)_0 \\
= \; & W^*(x, \mu, s) \\
= \; & \mu h_0(x, a_0^*, s) + \gamma^* h_1(x, a_0^*, s) + \beta \mathrm{E}\left[W^*(\ell(x, a_0^*, s'), \varphi(\mu, \gamma^*), s') | s\right]
\end{aligned}$$

where the inequality follows from the fact that shifting the policies one period back of the plan $\boldsymbol{a}^*$ is a feasible plan for the $\mathbf{PP}_{\varphi(\mu,\widetilde{\gamma})}$ problem with initial conditions (x_1^*, s') and that $W^*(x_1^*,\ \varphi(\mu,\ \widetilde{\gamma}),\ \ s')$ is the optimal value of $\mathbf{PP}_{\varphi(\mu,\widetilde{\gamma})}$, the second inequality follows because $(\boldsymbol{a}^*, \gamma^*)$ is a saddle point of $\mathbf{SPP}_\mu$ and the equalities follow from definitions, Theorem 2 and (**??**).

Therefore γ^* solves the min part of the right side of **SPFE**.

Therefore (a_0^*, γ^*) is a saddle point of the right side of **SPFE,** this implies the first equality in

$$\begin{aligned}
& \min_{\gamma\geq 0}\max_{a\in\mathcal{A}}\left\{\mu h_0(x,\ a,\ s)+\gamma h_1(x,\ a,\ s)+\beta\ \mathrm{E}\left[W^*(x',\ \mu',\ s')\right|s]\right\} \\
= \ & \mu h_0(x,\ a_0^*,\ s)+\gamma^* h_1(x,\ a_0^*,\ s)+\beta\ \mathrm{E}\left[W^*(x_1^*,\ \varphi(\mu,\ \gamma^*),\ s')\right|s] \\
= \ & W^*(x,\mu,\ s)
\end{aligned}$$

and the second equality comes, again, from (**??**). This proves that W^* satisfies **SPFE.**■

The argument used in the proof of Theorem 3 can be iterated a finite number of times to show the underlying recursive structure of the $\mathbf{PP}_\mu$ formulation. If $\mathbf{PP}_\mu$ has a unique solution $\{a_t^*\}_{t=0}^\infty$ at (x, s), then by Theorem 1 there is a $\mathbf{SPP}_\mu$ at (x, s) with solution $(\{a_t^*\}_{t=0}^\infty, \gamma^*)$, which in turn defines a $\mathbf{PP}_{\varphi(\mu,\gamma^*)}$ problem. As it has been seen in the proof of Theorem 3, $\{a_t^*\}_{t=1}^\infty$ solves $\mathbf{PP}_{\varphi(\mu,\gamma^*)}$ at $(\ell(x,\ a_0^*,\ s), s_1)$ and by Theorem 1 there is a γ_1^* such that $(\{a_t^*\}_{t=1}^\infty, \gamma_1^*)$ solves $\mathbf{SPP}_{\varphi(\mu,\gamma^*)}$ at $(\ell(x,\ a_0^*,\ s), s_1)$. In turn, $\{a_t^*\}_{t=2}^\infty$ solves $\mathbf{PP}_{\varphi^{(2)}(\mu,\gamma^*)}$ at $(\ell^{(2)}(x,\ a_0^*,\ s), s_1)$ where $\varphi^{(2)}(\mu,\ \gamma^*) \equiv \varphi(\varphi(\mu,\ \gamma^*),\ \gamma_1^*,\ s_1)$ and $\ell^{(2)}(x,\ a_0^*,\ s) \equiv \ell(\ell(x,\ a_0^*,\ s),\ a_1^*,\ s_1)$. Similarly, let $\varphi^{(n+1)}(\mu,\ \gamma^*) \equiv \varphi(\varphi^{(n)}(\mu,\ \gamma^*),\ \gamma_n^*,\ s_n)$, then by recursively applying the argument of the proof of Theorem 3 we obtain the following result.

Corollary 3.1. (Recursivity of $\mathbf{PP}_\mu$). If $\mathbf{PP}_\mu$ satisfies the assumptions of Theorem 1 and has a unique solution $\{a_t^*\}_{t=0}^\infty$ at (x, s), then, for any (t, x_t^*, s_t), $\left\{a_{t+j}^*\right\}_{j=0}^\infty$ is the solution to $\mathbf{PP}_{\varphi^{(t)}(\mu,\gamma^*)}$ at (x_t^*, s_t), where γ^* is the minimizer of $\mathbf{SPP}_\mu$ at (x, s).

The value function has some interesting properties that we would like to emphasize. First, notice that

$$\begin{aligned}
W^*(x,\mu,s) &= f_{(x,\mu,s)}(\boldsymbol{a}^*)+\gamma^* g(\boldsymbol{a}^*)_0 \\
&= f_{(x,\mu,s)}(\boldsymbol{a}^*) \\
&= \mathrm{E}_0\sum_{j=0}^{l}\sum_{t=0}^{N_j}\beta^t\mu^j h_0^j(x_t^*,\ a_t^*,\ s_t),
\end{aligned}$$

therefore, if $\{a_t^*\}_{t=0}^{\infty}$ at (x, s) is unquely defined, then W^* has a unique representation

$$\begin{aligned} W^*(x, \mu, s) &= \sum_{j=0}^{l} \mu^j \omega_j^*(x, \mu, s) \\ &= \mu\omega^*(x, \mu, s) \end{aligned}$$

where, for $j = 0, ...k, \omega_j(x, \mu, s) \equiv \mathrm{E}_0 \sum_{t=0}^{\infty} \beta^t h_0^j(x_t^*,\, a_t^*,\, s_t)$, and, for $j = k + 1, ...l, \omega_j(x, \mu, s) \equiv h_0^j(x_0^*,\, a_0^*,\, s_0)$. Similarly, the value function of $\mathbf{SPP}_{\varphi(\mu,\,\gamma^*)}$ at (x_1^*, s_1), $x_1^* = \ell(x,\, a_0^*,\, s)$, satisfies

$$W^*(x_1^*,\, \varphi(\mu,\, \gamma^*),\, s_1) \equiv \varphi(\mu,\, \gamma^*)\omega^*(x_1^*,\, \varphi(\mu,\, \gamma^*),\, s_1).$$

This representation not only has an interesting economic meaning – for example, as a 'social welfare function,' with varying weights, in problems with intertemporal participation constraints – but is also very convenient analytically. In particular, this reprensentation shows[10] that W^* is *convex and homogenous of degree one* in μ, with $W^*(x, 0, s) = 0$, for all (x, s)[11]. In addition, the following Corollary to Theorem 3 also shows that W^* satisfies what we call the *saddle-point inequality property* **SPI**. Lemmas 1 and 2 below show how these properties are extended to general W functions satisfying **SPFE**.

A function $W(x, \mu, s) = \sum_{j=0}^{l} \mu^j \omega_j(x, \mu, s)$ satisfies the *saddle-point inequality property* **SPI** at $(x,\, \mu,\, s)$ if and only if there exist (a^*, γ^*) satisfying

$$\begin{aligned} &\mu h_0(x,\, a^*,\, s) + \widetilde{\gamma} h_1(x,\, a^*,\, s) + \beta\, \mathrm{E}\left[\varphi(\mu,\, \widetilde{\gamma})\omega(x^{*\prime},\, \varphi(\mu,\, \gamma^*),\, s')|\, s\right] \\ &\geq \mu h_0(x,\, a^*,\, s) + \gamma^* h_1(x,\, a^*,\, s) + \beta\, \mathrm{E}\left[\varphi(\mu,\, \gamma^*)\omega(x^{*\prime},\, \varphi(\mu,\, \gamma^*),\, s')|\, s\right]. \quad (18) \\ &\geq \mu h_0(x,\, \widetilde{a},\, s) + \gamma^* h_1(x,\, \widetilde{a},\, s) + \beta\, \mathrm{E}\left[\varphi(\mu,\, \gamma^*)\omega(\widetilde{x}',\, \varphi(\mu,\, \gamma^*),\, s')|\, s\right], \quad (19) \end{aligned}$$

for any $\widetilde{\gamma} \in R_+^{l+1}$ and $(\widetilde{a}, \widetilde{x}')$ satisfying the technological constraints at $(x,\, s)$; that is, in **SPI** the multiplier minimization is taken in relation to the optimal continuation values.

Corollary 3.2. **($\mathbf{SPP}_\mu \Longrightarrow$SPI).** Let $W^*(x, \mu, s) \equiv V_\mu(x, s)$ be the value of $\mathbf{SPP}_\mu$ at (x, s), for an arbitrary $(x,\, \mu,\, s)$, then $W^*(x, \mu, s) = \sum_{j=0}^{l} \mu^j \omega_j^*(x, \mu, s)$ satisfies **SPI**.

Proof: We only need to show that (**??**) is satisfied, but this is immediate from the following identities

$$\begin{aligned} f_{(x,\mu,s)}(\boldsymbol{a}^*) &= \mu h_0(x,\, a_0^*,\, s) + \beta\, \mathrm{E}\left[\sum_{j=0}^{k} \mu^j \omega_j^*(x_1^*,\, \varphi(\mu,\, \gamma^*),\, s_1)|\, s\right] \\ \gamma g(\boldsymbol{a}^*)_0 &= \gamma\left[h_1(x,\, a_0^*,\, s) + \beta\, \mathrm{E}\left[\omega^*(x_1^*,\, \varphi(\mu,\, \gamma^*),\, s_1)|\, s\right]\right], \end{aligned}$$

[10]See Lemma 2A in the Appendix B.

[11]A function which is convex, homogeneous of degree one and finite at 0, is also called a *sublinear function* (see Rockafellar, 1981, p.29).

and the definition of $\mathbf{SPP}_\mu$ at (x,s); that is, for any $\widetilde{\gamma} \in R_+^{l+1}$,

$$\begin{aligned} & \mu h_0(x,\, a^*,\, s) + \widetilde{\gamma} h_1(x,\, a^*,\, s) + \beta \,\mathrm{E}\left[\varphi(\mu,\, \widetilde{\gamma})\omega^*(x^{*\prime},\, \varphi(\mu,\, \gamma^*),\, s')|\, s\right] \\ = \;& f_{(x,\mu,s)}(\boldsymbol{a}^*) + \widetilde{\gamma} g(\boldsymbol{a}^*)_0 \\ \geq \;& f_{(x,\mu,s)}(\boldsymbol{a}^*) + \gamma^* g(\boldsymbol{a}^*)_0 \\ = \;& \mu h_0(x,\, a^*,\, s) + \gamma^* h_1(x,\, a^*,\, s) + \beta \,\mathrm{E}\left[\varphi(\mu,\, \gamma^*)\omega^*(x^{*\prime},\, \varphi(\mu,\, \gamma^*),\, s')|\, s\right] \end{aligned}$$

■

We now show that, under fairly general conditions, programs satisfying **SPFE** are solutions to $\mathbf{SPP}_\mu$ at (x,s). More formally,

Theorem 4 (SPFE $\Longrightarrow$ $\mathbf{SPP}_\mu$) Assume W, satisfying **SPFE**, is continuous in (x,μ) and convex and homogeneous of degree one in μ. If the *SP policy correspondence* Ψ_W, associated with W, generates a solution $(\mathbf{a}^*, \boldsymbol{\gamma}^*)_{(x,\mu,s)}$, where $(\mathbf{a}^*)_{(x,\mu,s)}$ is uniquely determined, then $(\mathbf{a}^*, \boldsymbol{\gamma}^*)_{(x,\mu,s)}$ is also a solution to $\mathbf{SPP}_\mu$ at (x,s).

Notice that the assumptions on W are very general. In particular, if $W(x,\mu,s)$ is the value function of $\mathbf{SPP}_\mu$ at (x,s) (i.e. $W(x,\mu,s) \equiv V_\mu(x,s)$) then (as Lemma 2A in Appendix B shows) it is convex and homogeneous of degree one in μ and, if **A2** - **A5** are satisfied it is continuous and bounded in (x,μ). The only 'stringent condition' is that $(\mathbf{a}^*)_{(x,\mu,s)}$ must be uniquely determined, which is the case when W is concave in x and **A6s** is satisfied (see Corollary 4.1.)[12].

Before proving these results, we show that, as we have seen for W^*, convex and homogeneous functions W satisfying **SPFE** have some interesting properties, which are used in the proof of Theorem 4. First, without loss of generality (see **F2** and **F3** in Appendix C), we can express the *recursive equation* (**??**) in the form

$$\begin{aligned} \mu\omega^d(x,\mu,s) \;\; = \;\; & \mu h_0(x,\, a^*,\, s) + \gamma^* h_1(x,\, a^*,\, s) \\ & + \beta\,\mathrm{E}\left[\varphi(\mu,\, \gamma)\omega^{d'}(x^{*\prime},\, \varphi(\mu,\, \gamma^*),\, s')|\, s\right], \end{aligned} \tag{20}$$

where $\mu\omega^d(x,\mu,s) = W(x,\mu,s)$, and the vectors ω^d and $\omega^{d'}$are (partial) directional derivatives, in μ,of $W(x,\mu,s)$ and $W(x^{*\prime},\, \varphi(\mu,\, \gamma^*),\, s')$, respectively. Theredfore, the **SPFE** *saddle-point inequalities* take the form

$$\mu h_0(x,\, a^*,\, s) + \widetilde{\gamma} h_1(x,\, a^*,\, s) + \beta\,\mathrm{E}\left[\varphi(\mu,\, \widetilde{\gamma})\omega^{d'}(x^{*\prime},\, \varphi(\mu,\, \widetilde{\gamma}),\, s')|\, s\right]$$

[12]Marimon, Messner and Pavoni (2010) analyzes the general case with policy correspondences (multiple $a^*(x,\mu,s)$), using the subdiferential approach discussed here. We are specially grateful to the latter two authors who provided an example showing the problems that may arise when solutions are not uniquely determined. The proof of Theorem 4 has benefited from the study of their example. In parallel work, Cole and Kubler (2010) have recently provided an alternatve solution to the non-uniqueness case, based on 'lotteries'. Marimon and Werner (2011) furhter study the underlying envelope problem without differentiability.

$$\geq \mu h_0(x,\, a^*,\, s) + \gamma^* h_1(x,\, a^*,\, s) + \beta\ \mathrm{E}\left[\varphi(\mu,\, \gamma^*)\omega^{d'}(x^{*\prime},\, \varphi(\mu,\, \gamma^*),\, s')|\, s\right] \quad (21)$$

$$\geq \mu h_0(x, \widetilde{a},\, s) + \gamma^* h_1(x,\, \widetilde{a},\, s) + \beta\ \mathrm{E}\left[\varphi(\mu,\, \gamma^*)\omega^{d'}(\widetilde{x}',\, \varphi(\mu,\, \gamma^*),\, s')|\, s\right], \quad (22)$$

for any $\widetilde{\gamma} \in R_+^{l+1}$ and $(\widetilde{a}, \widetilde{x}')$ satisfying the technological constraints at $(x,\ s)$.

Second, as we show in Lemma 1, there is an equivalence between this **SPFE** property and the *saddle-point inequality property,* **SPI**, which substitutes (**??**) for

$$\mu h_0(x,\, a^*,\, s) + \widetilde{\gamma} h_1(x,\, a^*,\, s) + \beta\ \mathrm{E}\left[\varphi(\mu,\, \widetilde{\gamma})\omega^{d'}(x^{*\prime},\, \varphi(\mu,\, \gamma^*),\, s')|\, s\right]$$

$$\geq \mu h_0(x,\, a^*,\, s) + \gamma^* h_1(x,\, a^*,\, s) + \beta\ \mathrm{E}\left[\varphi(\mu,\, \gamma^*)\omega^{d'}(x^{*\prime},\, \varphi(\mu,\, \gamma^*),\, s')|\, s\right]. \quad (23)$$

Third, as we show in Lemma 2, if in addition $(\mathbf{a}^*, \boldsymbol{\gamma}^*)_{(x,\mu,s)}$ is uniquely determined, then W is differentiable in μ.

Lemma 1 (SPI $\Longleftrightarrow$ SPFE). If $W(x, \cdot, s)$ is convex and homogeneous of degree one, then (**??**) is satisfied if and only if (**??**) is satisfied. Furthermore, the inequality (**??**) is satisfied if and only if the following conditions are satisfied, for $j = 0, ..., l$,

$$\begin{aligned} h_1^j(x,\, a_0^*,\, s) + \beta\ \mathrm{E}\left[\omega_j^{d'}(x_1^*,\, \mu_1^*,\, s_1)|\, s\right] &\geq 0 && (24)\\ \gamma^{*j}\left[h_1^j(x,\, a_0^*,\, s) + \beta\ \mathrm{E}\left[\omega_j^{d'}(x_1^*,\, \mu_1^*,\, s_1)|\, s\right]\right] &= 0. && (25)\end{aligned}$$

Proof of Lemma 1: That **SPI** $\Longrightarrow$ **SPFE** follows from **F4** (see Appendix C). With respect to $W(x^{*\prime},\, \varphi(\mu,\, \gamma),\, s')$, **F4** takes the form:

$$\varphi(\mu,\, \widetilde{\gamma})\omega^{d'}(x^{*\prime},\, \varphi(\mu,\, \widetilde{\gamma}),\, s') \geq \varphi(\mu,\, \widetilde{\gamma})\omega^{d'}(x^{*\prime},\, \varphi(\mu,\, \gamma^*),\, s');$$

therefore (**??**) together with this latter inequality results in the following inequalities, which show that (**??**) is satisfied whenever (**??**) is satisfied:

$$\begin{aligned} &\mu h_0(x,\, a^*,\, s) + \widetilde{\gamma} h_1(x,\, a^*,\, s) + \beta\ \mathrm{E}\left[\varphi(\mu,\, \widetilde{\gamma})\omega^{d'}(x^{*\prime},\, \varphi(\mu,\, \widetilde{\gamma}),\, s')|\, s\right]\\ \geq\ &\mu h_0(x,\, a^*,\, s) + \widetilde{\gamma} h_1(x,\, a^*,\, s) + \beta\ \mathrm{E}\left[\varphi(\mu,\, \widetilde{\gamma})\omega^{d'}(x^{*\prime},\, \varphi(\mu,\, \gamma^*),\, s')|\, s\right]\\ \geq\ &\mu h_0(x,\, a^*,\, s) + \gamma^* h_1(x,\, a^*,\, s) + \beta\ \mathrm{E}\left[\varphi(\mu,\, \gamma^*)\omega^{d'}(x^{*\prime},\, \varphi(\mu,\, \gamma^*),\, s')|\, s\right]\end{aligned}$$

To see that **SPFE** $\Longrightarrow$ **SPI,** let

$$G_{(x,\, a^*,\, s)}(\gamma, \mu) \equiv \mu h_0(x,\, a^*,\, s) + \gamma h_1(x,\, a^*,\, s) + \beta\, \mathrm{E}\left[\varphi(\mu,\, \gamma)\omega^{d'}(x^{*\prime},\, \varphi(\mu,\, \gamma),\, s')|\, s\right],$$

and

$$F_{(x,\, a^*,\, s)}(\gamma, \mu) \equiv \mu h_0(x,\, a^*,\, s) + \gamma h_1(x,\, a^*,\, s) + \beta\, \mathrm{E}\left[\varphi(\mu,\, \gamma)\omega^{d'}(x^{*\prime},\, \varphi(\mu,\, \gamma^*),\, s')|\, s\right],$$

Then (**??**) reduces to $G_{(x,\, a^*,\, s)}(\gamma, \mu) \geq G_{(x,\, a^*,\, s)}(\gamma^*, \mu)$ and (**??**) to $F_{(x,\, a^*,\, s)}(\gamma, \mu) \geq F_{(x,\, a^*,\, s)}(\gamma^*, \mu)$. Since $G_{(x,\, a^*,\, s)}(\gamma^*, \mu) = F_{(x,\, a^*,\, s)}(\gamma^*, \mu)$, the above inequalities (**??**) show that, if $f_{(x,\, a^*,\, s)}(\gamma^*, \mu) \in \partial_\gamma F_{(x,\, a^*,\, s)}(\gamma^*, \mu)$, for all $\gamma \geq 0$, then

$$\begin{aligned} G_{(x,\, a^*,\, s)}(\gamma, \mu) - G_{(x,\, a^*,\, s)}(\gamma^*, \mu) \quad &\geq \quad F_{(x,\, a^*,\, s)}(\gamma, \mu) - F_{(x,\, a^*,\, s)}(\gamma^*, \mu) \\ &\qquad (\gamma - \gamma^*)\, f_{(x,\, a^*,\, s)}(\gamma^*, \mu); \end{aligned}$$

that is, $f_{(x,\, a^*,\, s)}(\gamma^*, \mu) \in \partial_\gamma G_{(x,\, a^*,\, s)}(\gamma^*, \mu)$.

Let $g_{(x,\, a^*,\, s)}(\gamma^*, \mu)$ be an exreme point of $\partial_\gamma G_{(x,\, a^*,\, s)}(\gamma^*, \mu)$, since $G_{(x,\, a^*,\, s)}(\gamma, \mu)$ is homogenous of degree one in γ, by **F2** $\exists\ \gamma_k \longrightarrow \gamma^*$with G differentiable at γ_k and $\nabla G_{(x,\, a^*,\, s)}(y_k, \mu) \longrightarrow g_{(x,\, a^*,\, s)}(\gamma^*, \mu)$. By homogeneity of degree zero of $\omega^{d'}(x^{*\prime},\, \mu',\, s')$ with respect to μ'

$$\nabla G_{(x,\, a^*,\, s)}(\gamma_k, \mu) = h_1(x,\, a^*(x,\, \mu,\, s),\, s) + \beta\, \mathrm{E}\left[\omega^{d'}(x^{*\prime}(x,\, \mu,\, s),\, \varphi(\, \mu, \gamma_k),\, s') |\, s\right].$$

Given the differentiability of $\nabla G_{(x,\, a^*,\, s)}(y_k, \mu)$ at γ_k, continuity of φ and $\omega^{d'}$, implies that

$$g_{(x,\, a^*,\, s)}(\gamma^*, \mu) = h_1(x,\, a^*(x,\, \mu,\, s),\, s) + \beta\, \mathrm{E}\left[\omega^{d'}(x^{*\prime}(x,\, \mu,\, s),\, \varphi(\, \mu, \gamma^*),\, s') |\, s\right],$$

and, therefore, $g_{(x,\, a^*,\, s)}(\gamma^*, \mu) \in \partial_\gamma F_{(x,\, a^*,\, s)}(\gamma^*, \mu)$ – in fact, it is also an extreme point of $\partial_\gamma F_{(x,\, a^*,\, s)}(\gamma^*, \mu)$. This shows that $\partial_\gamma F_{(x,\, a^*,\, s)}(\gamma^*, \mu) = \partial_\gamma G_{(x,\, a^*,\, s)}(\gamma^*, \mu)$, which, in turn, implies the equivalence between (**??**) and (**??**).

Finally, the proof of the Kuhn-Tucker conditions is standard. First, necessity of (**??**) follows from the fact that $\gamma^* \geq 0$ is finite, which will not be the case if, for some $j = 0, ..., l$,

$$h_1^j(x,\, a_0^*,\, s) + \beta\ \mathrm{E}\left[\omega_j^{d'}(x_1^*,\, \mu_1^*,\, s_1) |\, s\right] < 0.$$

To see necessity of (**??**), let $\gamma_{(i)}^{*j} = \gamma^{*j}$, if $j \neq i$, and $\gamma_{(i)}^{*i} = 0$, then (**??**) results in:

$$\mu h_0(x,\, a^*,\, s) + \gamma_{(i)}^* h_1(x,\, a^*,\, s) + \beta\ \mathrm{E}\left[\varphi(\mu,\, \gamma_{(i)}^*)\omega^{d'}(x^{*\prime},\, \varphi(\mu,\, \gamma^*),\, s') |\, s\right]$$

$$\geq \mu h_0(x,\, a^*,\, s) + \gamma^* h_1(x,\, a^*,\, s) + \beta\ \mathrm{E}\left[\varphi(\mu,\, \gamma^*)\omega^{d'}(x^{*\prime},\, \varphi(\mu,\, \gamma^*),\, s') |\, s\right],$$

which, together with (**??**), implies that

$$0 \geq \gamma^{*j}\left[h_1^j(x,\, a_0^*,\, s) + \beta\ \mathrm{E}\left[\omega_j^{d'}(x_1^*,\, \mu_1^*,\, s_1) |\, s\right]\right] \geq 0.$$

To see, that (**??**) and (**??**) imply (**??**), suppose they are satisfied and there exists a $\widetilde{\gamma} \geq 0$, for which (**??**) it is not, then it must be that

$$\begin{aligned} & \widetilde{\gamma}\left[h_1(x,\, a_0^*,\, s) + \beta\ \mathrm{E}\left[\omega^{d'}(x_1^*,\, \mu_1^*,\, s_1)|\, s\right]\right] \\ < \quad & \gamma^*\left[h_1(x,\, a_0^*,\, s) + \beta\ \mathrm{E}\left[\omega^{d'}(x_1^*,\, \mu_1^*,\, s_1)|\, s\right]\right] = 0, \end{aligned}$$

which contradicts (**??**)■

Lemma 2. If $(\mathbf{a}^*, \boldsymbol{\gamma}^*)_{(x,\mu,s)}$ is generated by $\Psi_W(x,\, \mu,\, s)$ and $(\mathbf{a}^*)_{(x,\mu,s)}$ is uniquely defined then $W(x_t^*, \mu_t^*, s_t)$ is differentiable with respect to μ_t^*; for every (x_t^*, μ_t^*, s_t),with (x_t^*, μ_t^*) realized by[13] $(\mathbf{a}^*, \boldsymbol{\gamma}^*)_{(x,\mu,s)}$.

Proof of Lemma 2: By (**??**) the recursive equation (**??**) simplifies to

$$\mu\omega^d(x, \mu, s) = \mu h_0(x,\, a^*,\, s) + \beta\ \mathrm{E}\left[\sum_{j=0}^{k} \mu^j \omega_j^{d'}(x^{*\prime},\, \mu + \gamma^*),\, s')|\, s\right],$$

Assume, for the moment, that $(\mathbf{a}^*, \boldsymbol{\gamma}^*)_{(x,\mu,s)}$ is uniquely determined. By recursive iteration, it follows that

$$\begin{aligned} \mu\omega^d(x_0, \mu_0, s_0) &= \mu h_0(x_0,\, a_0^*,\, s_0) \\ &\quad +\beta\ \mathrm{E}_0\left[\sum_{j=0}^{k} \mu^j\left(h_0^j(x_1^*,\, a_1^*,\, s_1) + \beta\omega_j^{d''}(x_2^*,\, \mu + \gamma_0^* + \gamma_1^*),\, s_2)\right)|\, s_0\right] \\ &= \mu h_0(x_0,\, a_0^*,\, s_0) + \beta\ \mathrm{E}_0\left[\sum_{j=0}^{k}\mu^j \sum_{t=1}^{\infty} \beta^t h_0^j(x_t^*,\, a_t^*,\, s_t)|\, s_0\right]; \end{aligned}$$

therefore, uniqueness of $(\mathbf{a}^*, \boldsymbol{\gamma}^*)_{(x,\mu,s)}$ implies: *i)* $\omega^d(x,\, \mu,\, s)$ is uniquely defined: $\omega^d(x,\, \mu,\, s) = \omega(x,\, \mu,\, s) \equiv \nabla_\mu W(x,\, \mu,\, s)$; which, in turn, implies that $W(x, \cdot, s)$ is differentiable, and *ii)* $\omega_j(x,\, \mu,\, s) = \mathrm{E}_0 \sum_{t=0}^{\infty} \beta^t h_0^j(x_t^*,\ a^*(x_t^*,\, \mu_t^*,\, s_t),\, s_t)$, for $j = 0, \ldots k$ (with $(x_0^*,\, \mu_0^*,\, s_0) \equiv (x,\, \mu,\, s)$, $x_{t+1}^* = \ell(x_t^*,\, a^*(x_t^*,\, \mu_t^*,\, s_t),\, s_t)$, and $\mu_{t+1}^* = \mu^* + \gamma^*(x_t^*,\, \mu_t^*,\, s_t)$), and $\omega_j(x,\, \mu,\, s) = h_0^j(x,\, a^*(x,\, \mu,\, s),\, s)$, for $j = k+1 \ldots l$

Given $(\mathbf{a}^*)_{(x,\mu,s)}$, supose now $\left(\mathbf{a}^*, \widetilde{\boldsymbol{\gamma}}^*\right)_{(x,\mu,s)}$ is also generated by $\Psi_W(x,\, \mu,\, s)$. Both saddle-point paths must have the same value (see Lemma 3A in

[13]That is, $(x_0^*, \mu_0^*) \equiv (x, \mu), x_{t+1}^* = \ell(x_t^*, a_t^*, s_{t+1})$ and $\mu_{t+1}^* = \varphi(\mu_t^*, \gamma_t^*)$.

Appendix B); in particular, following the same recursive argument:

$$\begin{aligned}\mu\omega^d(x_0,\mu_0,s_0) &= \mu h_0(x_0,\, a_0^*,\, s_0) + \beta\,\mathrm{E}\left[\sum_{j=0}^{k}\mu^j\omega_j^{d'}(x^{*\prime},\,\mu+\widetilde{\gamma}^*),\, s')|\, s\right] \\ &= \mu h_0(x_0,\, a_0^*,\, s_0) + \beta\,\mathrm{E}_0\left[\sum_{j=0}^{k}\mu^j\sum_{t=1}^{\infty}\beta^t h_0^j(x_t^*,\, a_t^*,\, s_t)|\, s_0\right],\end{aligned}$$

which proves the differentiablity of W with respect to μ, even when $(\boldsymbol{\gamma}^*)_{(x,\mu,s)}$ is not uniquely determined (i.e. there may be kinks in the Pareto frontier)■

An immediate, and important, consequence of Lemma 2 is the following result:

Corollary: If $(\mathbf{a}^*)_{(x,\mu,s)}$ is uniquely defined by $\Psi_W(x,\,\mu,\,s)$, from any initial condition (x,μ,s), then the following (recursive) equations are satisfied:

$$\omega_j(x,\,\mu,\,s) = h_0^j(x,\,a^*(x,\,\mu,\,s),\,s) + \beta\,\mathrm{E}\left[\omega_j(x^{*\prime}(x,\,\mu,\,s),\,\mu^{*\prime}(x,\,\mu,\,s),\,s')|\,s\right], \quad \text{if} \quad j=0,...,k, \quad \text{and} \tag{26}$$

$$\omega_j(x,\,\mu,\,s) = h_0^j(x,\,a^*(x,\,\mu,\,s),\,s) \quad \text{if} \quad j=k+1,...,l. \tag{27}$$

Furtheremore, $(\mathbf{a}^*)_{(x,\mu,s)}$ is uniquely defined by $\Psi_W(x,\,\mu,\,s)$ whenever $W(\cdot,\,\mu,\,s)$ is concave and **A6s** is satisfied.

Notice that, in proving Lemma 2, uniqueness of the solution paths has implied uniqueness of the value function decomposition: $W = \mu\omega$. This unique decomposition has implied the recursive equations (**??**) and (**??**). Uniqueness of the value function decomposition is equivalent to the differentiability of the value function. In fact, once it has been established that the value function is differentiable, one can obtain equations (**??**) and (**??**) as a simple application of the *Envelope Theorem*. For example, equation (**??**) is just[14]:

$$\partial_j W(x,\,\mu,\,s) = h_0^j(x,\,a^*(x,\,\mu,\,s),\,s) + \beta\,\mathrm{E}\left[\partial_j W(x^{*\prime}(x,\,\mu,\,s),\,\mu^{*\prime}(x,\,\mu,\,s),\,s')|\,s\right].$$

We now turn to the proof of Theorem 4, where these recursive equations (**??**) and (**??**) play a key role.

Proof (Theorem 4): By Lemma 2, there is a unique representation $W(x,\,\mu,\,s) = \mu\omega(x,\,\mu,\,s)$. To see that solutions of **SPFE** satisfy the participation constraints of $\mathbf{SPP}_\mu$, we use the first order conditions (**??**) and **??**), as well as the individual recursive equations (**??**) and **??**) of the previous Corollary. As in the proof of Lemma 2, equation (**??**) can be iterated to obtain

$$\omega_j(x,\,\mu,\,s) = \mathrm{E}_0\left[\sum_{t=0}^{\infty}\beta^t h_0^j(x_t^*,\,a_t^*,\,s_t)|\,s\right], \text{ if } j=0,...,k. \tag{28}$$

[14] We use the standard notation $\partial_j W(x,\,\mu,\,s) \equiv \frac{\partial W(x,\mu,s)}{\partial \mu_i}$, and also $\omega_j(x,\,\mu,\,s) \equiv \partial_j W(x,\,\mu,\,s)$

Following the same steps for any $t > 0$ and state $(x_t^*,\, \mu_t^*,\, s_t)$, equation(**??**) and (**??**) together with the inequality (**??**) show that the intertemporal participation constraints in $\mathbf{PP}_\mu$ – and therefore in $\mathbf{SPP}_\mu$ – are satisfied; that is,

$$\mathrm{E}_t \sum_{n=1}^{N_j+1} \beta^n \; h_0^j(x_{t+n}^*,\, a_{t+n}^*,\, s_{t+n}) + h_1^j(x_t^*,\, a_t^*,\, s_t) \geq 0, \; ; \; t \geq 0, \quad j = 0, ..., l \tag{29}$$

Now, to see that solutions of **SPFE** are, in fact, solutions of $\mathbf{SPP}_\mu$ we argue by contradiction. Suppose there exist a program $\{\widetilde{a}_t\}_{t=0}^\infty$, and $\{\widetilde{x}_t\}_{t=0}^\infty$, $\widetilde{x}_0 = x, \widetilde{x}_{t+1} = \ell(\widetilde{x}_t,\, \widetilde{a}_t,\, s_{t+1})$, satisfying the constraints of $\mathbf{SPP}_\mu$ with initial condition (x, s) and such that

$$\begin{aligned} & \mu h_0(x,\, \widetilde{a}_0,\, s) + \gamma^* h_1(x,\, \widetilde{a}_0,\, s) \\ + & \beta \mathrm{E} \left[\sum_{j=0}^{k} \left(\mu^j + \gamma^{*j}\right) \sum_{n=1}^{\infty} \beta^t \; h_0^j(\widetilde{x}_t,\, \widetilde{a}_t,\, s_t) + \sum_{j=k+1}^{l} \gamma^{*j} h_0^j(\widetilde{x}_1,\, \widetilde{a}_1,\, s_1) |\, s \right] \\ > \; & \mu h_0(x,\, a_0^*,\, s) + \gamma^* h_1(x,\, a_0^*,\, s) \\ + & \beta \mathrm{E} \left[\sum_{j=0}^{k} \left(\mu^j + \gamma^{*j}\right) \sum_{n=1}^{\infty} \beta^t \; h_0^j(x_t^*,\, a_t^*,\, s_t) + \sum_{j=k+1}^{l} \gamma^{*j} h_0^j(x_1^*,\, a_1^*,\, s_1) |\, s \right]. \end{aligned} \tag{30}$$

The following string of equalities and inequalities, which we explain at the end, contradict this inequality,

$$
\begin{aligned}
&\mu h_0(x,\, a_0^*,\, s) + \gamma_0^* h_1(x,\, a_0^*,\, s) \\
&+ \beta \mathrm{E}\left[\sum_{j=0}^{k} \left(\mu^j + \gamma^{*j}\right) \sum_{n=1}^{\infty} \beta^t \; h_0^j(x_t^*,\, a_t^*,\, s_t) + \sum_{j=k+1}^{l} \gamma^{*j} h_0^j(x_1^*,\, a_1^*,\, s_1) |\, s\right] \\
&= \mu h_0(x,\, a_0^*,\, s) + \gamma_0^* h_1(x,\, a_0^*,\, s) + \beta\, \mathrm{E}\left[\mu_1^* \omega(x_1^*,\, \mu_1^*,\, s_1) |\, s\right] && (31) \\
&\geq \mu h_0(x,\, \widetilde{a}_0,\, s) + \gamma_0^* h_1(x,\, \widetilde{a}_0,\, s) + \beta\, \mathrm{E}\left[\mu_1^* \omega(\widetilde{x}_1,\, \mu_1^*,\, s_1) |\, s\right] && (32) \\
&= \mu h_0(x,\, \widetilde{a}_0,\, s) + \gamma_0^* h_1(x,\, \widetilde{a}_0,\, s) \\
&+ \beta\, \mathrm{E}\left[\mu_1^* h_0(\widetilde{x}_1,\, a^*(\widetilde{x}_1,\, \mu_1^*, s_1),\, s_1) + \gamma^*(\widetilde{x}_1,\, \mu_1^*, s_1) h_1(\widetilde{x}_1,\, a^*(\widetilde{x}_1,\, \mu_1^*, s_1),\, s_1)\right. && (33) \\
&+ \left.\beta \mu^{*\prime}(\widetilde{x}_1,\, \mu_1^*, s_1) \omega(x^{*\prime}(\widetilde{x}_1,\, \mu_1^*, s_1),\, \mu^{*\prime}(\widetilde{x}_1,\, \mu_1^*, s_1), s_2) |\, s\right] \\
&\geq \mu h_0(x,\, \widetilde{a}_0,\, s) + \gamma_0^* h_1(x,\, \widetilde{a}_0,\, s) \\
&+ \beta\, \mathrm{E}\left[\mu_1^* h_0(\widetilde{x}_1,\, \widetilde{a}_1,\, s_1) + \gamma^*(\widetilde{x}_1,\, \mu_1^*, s_1) h_1(\widetilde{x}_1,\, \widetilde{a}_1,\, s_1)\right. && (34) \\
&+ \left.\beta \mu^{*\prime}(\widetilde{x}_1,\, \mu_1^*, s_1) \omega(\widetilde{x}_2,\, \mu^{*\prime}(\widetilde{x}_1,\, \mu_1^*, s_1), s_2) |\, s\right] \\
&\geq \mu h_0(x,\, \widetilde{a}_0,\, s) + \gamma_0^* h_1(x,\, \widetilde{a}_0,\, s) \\
&+ \beta\, \mathrm{E}\left[\mu_1^* h_0(\widetilde{x}_1,\, \widetilde{a}_1,\, s_1) + \gamma^*(\widetilde{x}_1,\, \mu_1^*, s_1) h_1(\widetilde{x}_1,\, \widetilde{a}_1,\, s_1) + \beta \mu^{*\prime}(\widetilde{x}_1,\, \mu_1^*, s_1) \omega(\widetilde{x}_2,\, \mu_1^*, s_2) |\, s\right] && (35) \\
&\geq \mu h_0(x,\, \widetilde{a}_0,\, s) + \gamma_0^* h_1(x,\, \widetilde{a}_0,\, s) \\
&+ \beta\, \mathrm{E}\left[\mu_1^* \left[h_0(\widetilde{x}_1,\, \widetilde{a}_1,\, s_1) + \beta \omega(\widetilde{x}_2,\, \mu_1^*, s_2)\right] |\, s\right] && (36) \\
&= \mu h_0(x,\, \widetilde{a}_0,\, s) + \gamma_0^* h_1(x,\, \widetilde{a}_0,\, s) + \beta\, \mathrm{E}\left[\mu_1^* h_0(\widetilde{x}_1,\, \widetilde{a}_1,\, s_1) |\, s\right] \\
&+ \beta^2\, \mathrm{E}\left[\mu_1^* h_0(\widetilde{x}_2,\, a^*(\widetilde{x}_2,\, \mu_1^*, s_2),\, s_2) + \gamma^*(\widetilde{x}_2,\, \mu_1^*, s_2) h_1(\widetilde{x}_2,\, a^*(\widetilde{x}_2,\, \mu_1^*, s_2),\, s_2)\right. && (37) \\
&+ \left.\beta \mu^{*\prime}(\widetilde{x}_2,\, \mu_1^*, s_2) \omega(x^{*\prime}(\widetilde{x}_2,\, \mu_1^*, s_2),\, \mu^{*\prime}(\widetilde{x}_2,\, \mu_1^*, s_2), s_2) |\, s\right] \\
&\geq \mu h_0(x,\, \widetilde{a}_0,\, s) + \gamma_0^* h_1(x,\, \widetilde{a}_0,\, s) \\
&+ \beta \mathrm{E}\left[\sum_{j=0}^{k} \left(\mu^j + \gamma^{*j}\right) \sum_{t=1}^{\infty} \beta^t \; h_0^j(\widetilde{x}_t; \widetilde{a}_t,\, s_t) + \sum_{j=k+1}^{l} \gamma^{*j} h_0^j(\widetilde{x}_1,\, \widetilde{a}_1,\, s_1) |\, s\right] && (38)
\end{aligned}
$$

Notice that the first equality (**??**) is just uses the value function decomposition, the other two equalities (**??**) and (**??**) are simple expansions of the of the saddle-point value paths (i.e., of (**??**)) and in these expansions equations (**??**) and (**??**) play a key role. Inequalities (**??**) and (**??**) follow from the maximality property of **SPFE.** Inequalities (**??**) and (**??**) require explanation. Inequality (**??**) follows from une of the properties of convex and homogeneous of degree one functions (i.e. **F4**: $\widehat{\mu}\omega(\widehat{\mu}) \geq \widehat{\mu}\omega(\mu)$, see Appendix), given that (**??**) is simply $\mu^{*\prime}(\widetilde{x}_1,\, \mu_1^*, s_1)\omega(\widetilde{x}_2,\, \mu^{*\prime}(\widetilde{x}_1,\, \mu_1^*, s_1) \geq \mu^{*\prime}(\widetilde{x}_1,\, \mu_1^*, s_1)\omega(\widetilde{x}_2,\, \mu_1^*, s_2)$. Inequality (**??**) follows from applying the slackness inequality (**??**), as well as equations (**??**) and (**??**) to the plan generated by **SPFE** in state $(\widetilde{x}_2,\, \mu_1^*, s_2)$ (i.e. to $\{a_t^*(\widetilde{x}_2,\, \mu_1^*, s_2)\}_{t=2}^{\infty}$); these inequalities are needed to show that such plan satisfies the corresponding **SPP** constraints (**??**); that is, $\left[h_1^j(\widetilde{x}_1,\, \widetilde{a}_1,\, s_1) + \beta\omega_j(\widetilde{x}_2,\, \mu_1^*, s_2)\right] \geq$

$0, j=0,...,l$. Finally, since the equality (**??**) is simply the equality (**??**) after one iteration, repeated iterations result in the last inequality (**??**), which contradicts (**??**).

It only remains to show that the inf part of **SPP** is also satisfied. Reasoning again by contradiction, suppose there exist a $\widetilde{\gamma} \geq 0$ such that

$$\begin{aligned}
& \mu h_0(x,\, a_0^*,\, s) + \widetilde{\gamma} h_1(x,\, a_0^*,\, s) \\
& + \beta \mathrm{E}\left[\sum_{j=0}^{k}\left(\mu^j + \widetilde{\gamma}^j\right)\sum_{n=1}^{\infty}\beta^t\; h_0^j(x_t^*,\, a_t^*,\, s_t) + \sum_{j=k+1}^{l}\widetilde{\gamma}^j h_0^j(x_1^*,\, a_1^*,\, s_1)|\, s\right] \\
& <\;\; \mu h_0(x,\, a_0^*,\, s) + \gamma^* h_1(x,\, a_0^*,\, s) \\
& + \beta \mathrm{E}\left[\sum_{j=0}^{k}\left(\mu^j + \gamma^{*j}\right)\sum_{n=1}^{\infty}\beta^t\; h_0^j(x_t^*,\, a_t^*,\, s_t) + \sum_{j=k+1}^{l}\gamma^{*j} h_0^j(x_1^*,\, a_1^*,\, s_1)|\, s\right].
\end{aligned} \tag{39}$$

Using the value function decomposition representation, this inequality can also be expressed as

$$\begin{aligned}
& \widetilde{\gamma}\left[h_1(x,\, a^*(x,\, \mu,\, s),\, s) + \beta\ \mathrm{E}\left[\omega(x^{*\prime}(x,\, \mu,\, s),\, \mu^{*\prime}(x,\, \mu,\, s),\, s')|\, s\right]\right] \\
< \;\; & \gamma^*(x,\, \mu,\, s)\left[h_1(x,\, a^*(x,\, \mu,\, s),\, s) + \beta\ \mathrm{E}\left[\omega_j(x^{*\prime}(x,\, \mu,\, s),\, \mu^{*\prime}(x,\, \mu,\, s),\, s')|\, s\right]\right],
\end{aligned}$$

but the first order conditions (**??**) and (**??**) require that (**??**) is satisfied, i.e.

$$\begin{aligned}
& \widetilde{\gamma}\left[h_1(x,\, a^*(x,\, \mu,\, s),\, s) + \beta\ \mathrm{E}\left[\omega(x^{*\prime}(x,\, \mu,\, s),\, \mu^{*\prime}(x,\, \mu,\, s),\, s')|\, s\right]\right] \\
\geq \;\; & \gamma^*(x,\, \mu,\, s)\left[h_1(x,\, a^*(x,\, \mu,\, s),\, s) + \beta\ \mathrm{E}\left[\omega_j(x^{*\prime}(x,\, \mu,\, s),\, \mu^{*\prime}(x,\, \mu,\, s),\, s')|\, s\right]\right] = 0
\end{aligned}$$

which contradicts (**??**)■

Corollary to Lemma 2 implies the followng Corollary to Theorem 4:

Corollary 4.1. Assume W, satisfying **SPFE,** is continuous in (x, μ), convex and homogeneous of degree one in μ, concave in x and that **A6s** is satisfied**.** If $(\mathbf{a}^*, \boldsymbol{\gamma}^*)_{(x,\mu,s)}$ is generated by $\Psi_W(x,\, \mu,\, s)$ then $(\mathbf{a}^*, \boldsymbol{\gamma}^*)_{(x,\mu,s)}$ is also a solution to $\mathbf{SPP}_\mu$ at (x, s).

5 DSPP and the contraction mapping

In this Section we show how our main results –Theorems 3 and 4– can also be obtained by applying the *Contraction Mapping Theorem* to the *Dynamic Saddle-Point Problem,* corresponding to **SPFE**. This Section provides more general sufficient conditions to obtain a solution to the original problem $\mathbf{PP}_\mu$ starting from **SPFE**. While these conditions are satisfied whenever the conditions of Theorem 4 are satisfied, they help to better understand the passage $\mathbf{SPFE}\rightarrow\mathbf{PP}_\mu$ and, in particular, they show how the standard method of value function iteration extends to our *saddle-point problems* and, therefore, that computing solutions to our original $\mathbf{PP}_\mu$ does not require special computational techniques. Furthermore, it also shows the interest of using the $W = \mu\omega$ representation in computing recursive contracts (i.e. taking ω as the starting vector valued function) and how, in contast with the 'APG approach' to solving contractual problems, 'promised values' are not part of the constraints ('promise keeping constraints'), but an outcome of the recursive contract.

We first define some spaces of "value" functions

$$\begin{aligned}\mathcal{M}_b =\{&W : X \times \mathcal{R}_+^{l+1} \times S \to \mathcal{R} \\ &i) W(\cdot,\, \cdot,\, s) \text{ is continuous, and } W(\cdot,\, \mu,\, s) \text{ bounded, when } \|\mu\| \leq 1, \\ &ii) W(x,\, \cdot,\, s) \text{ is convex and homogeneous of degree one}\}\end{aligned}$$

and

$$\begin{aligned}\mathcal{M}_{bc} =\{&W \in \mathcal{M}_b \text{ and} \\ &iii) W(\cdot,\, \mu,\, s) \text{ is concave}\}.\end{aligned}$$

$\mathcal{M}_b$ is a space of continuous, bounded functions (in x), and convex and homogenous of degree one (in μ)[15], while $\mathcal{M}_{bc}$ is the subspace of concave functions (in x). Both spaces are normed vector spaces with the norm

$$\|W\| = \sup\left\{|W(x, \mu, s)| : \|\mu\| \leq 1, x \in X, s \in S\right\}.$$

We show in Appendix D (Lemma 6A) that they are complete metric spaces; therefore, suitable spaces for the *Contraction Mapping Theorem.*

Since, whenever W satisfies (ii) it can be represented as $W(x,\, \mu,\, s) = \mu\omega(x,\, \cdot,\, s)$ (see Lemma 4A), it is convenient to define the corresponding spaces of functions:

$$\begin{aligned}M_b =\{&\omega : X \times \mathcal{R}_+^{l+1} \times S \to \mathcal{R}^{l+1} \text{ s.t., for } j = 0, ..., l, \\ &i) \omega_j(\cdot,\, \cdot,\, s) \text{ is continuous, and } \omega_j(\cdot,\, \mu,\, s) \text{ bounded, when } \|\mu\| \leq 1 \\ &ii) \omega_j(x,\, \cdot,\, s) \text{ is convex and homogeneous of degree zero}\}\end{aligned}$$

and

$$\begin{aligned}M_{bc} =\{&\omega \in M_b \text{ s.t., for } j = 0, ..., l, \\ &iii) \omega_j(\cdot,\, \mu,\, s) \text{ is concave}\}.\end{aligned}$$

Notice that $\omega \in M$ uniquely defines a function $W \in \mathcal{M}$, given by $W \equiv \mu\omega$, but $W \in \mathcal{M}$ does not uniquely define a $\mathcal{R}^{l+1}$ valued function $\omega \in M$; it does when, in addition, W is differentiable in μ (see Appendix C)[16].

---

[15] Without loss of generality, we could also require that $W(x, 0, s) < \infty$ and then replace (ii) by $W(x,\, \cdot,\, s)$ is *sublinear* (see footnote 12).

[16] M denotes either M_b or M_{bc}.

As we have seen in Section 4[17], when $W^{*}(x,\mu,s) = V_{\mu}(x,s)$ is the value of $\mathbf{SPP}_{\mu}$, with initial conditions (x,s), then $W^{*}(x,\mu,s) = \sum_{j=0}^{l}\mu^{j}\omega_{j}^{*}(x,\mu,s)$ with $W^{*} \in \mathcal{M}_{b}$, whenever **A2 - A4** are satisfied (and $W^{*} \in \mathcal{M}_{bc}$ if in addition **A5 - A6** are satisfied); furthermore, $\omega^{*} \in M$ is unique whenever $(\mathbf{a}^{*})_{(x,\mu,s)}$ is uniquely defined.

Given a function $\omega \in M$, and an initial condition (x,μ,s), we can define the following *Dynamic Saddle Point Problem*:

DSPP

$$\begin{aligned}&\inf_{\gamma\geq 0}\sup_{a}\left\{\mu h_{0}(x,\,a,\,s)+\gamma h_{1}(x,\,a,\,s)+\beta\,\mathrm{E}\left[\mu'\omega(x',\,\mu',\,s')|\,s\right]\right\}\\ &\text{s.t. } x'=\ell(x,\,a,\,s),\; p(x,a,\,s)\geq 0\\ &\text{and } \mu'=\varphi(\mu,\,\gamma),\end{aligned}$$

To guarantee that this problem has well defined solutions we make an interiority assumption:

A7b. For any $(x,s) \in X \times S$, there exists an $\widetilde{a} \in A$, satisfying $p(x,a,\,s) > 0$, such that, for any $\mu' \in R_{+}^{l+1}, \|\mu'\| < +\infty$, and $j = 0,...,l$: $h_{1}^{j}(x,\widetilde{a},s) + \beta\mathrm{E}\left[\omega^{j}(\ell(x,\,\widetilde{a},\,s'),\mu',s')|\,s\right] > 0$.

Notice that **A7b** is satisfied, whenever **A7** is satisfied and $\mu'\omega(\ell(x,\,\widetilde{a},\,s'),\mu',s')$ is the value function of $\mathbf{SPP}_{(\ell(x,\widetilde{a},s'),\mu',s')}$. In general, **A7b** is not a restrictive assumption in the class of possible value functions if the original problem has interior solutions. Nevertheless, an assumption, such as **A7**b is needed when one takes $\mathbf{DSPP}_{(x,\mu,s)}$ as the starting problem. This is a relatively standard min max problem, except for the dependency of ω on $\varphi(\mu,\,\gamma)$. The following proposition shows that it has a solution. Obviously, solutions to $\mathbf{DSPP}_{(x,\mu,s)}$ satisfy **SPFE.** . An immediate consequence of **A7b**, is the following lemma:

Lemma 3. Assume **A4** and **A7b** and let $\omega \in M_{b}$. There exist a $B > 0$, such that if $(a^{*}(x,\mu,s),\gamma^{*}(x,\mu,s))$ is a solution to **DSPP** at (x,μ,s), then $\|\gamma^{*}(x,\mu,s)\| \leq B\,\|\mu\|$.

Proof: Denote by $(a^{*},\gamma^{*}) \equiv (a^{*}(x,\mu,s),\gamma^{*}(x,\mu,s))$ the solution to **DSPP** at

[17] See also Lemma 2A in Appendix B.

(x, μ, s), and let $\widetilde{a}$ be the interior solution of **A7b**, then

$$\begin{aligned}
& \mu h_0(x,\, a^*,\, s) + \gamma^* h_1(x,\, a^*,\, s) + \beta\ \mathrm{E}\left[\mu'\omega(\ell(x,\, a^*, s),\, \varphi(\mu,\, \gamma^*),\, s')\right| s] \\
= \;& \mu h_0(x,\, a^*,\, s) + \beta\ \mathrm{E}\left[\sum_{j=0}^{k} \mu^j \omega_j(\ell(x,\, a^*, s),\, \varphi(\mu,\, \gamma^*),\, s')\right| s\Bigg] \\
+ \;& \gamma^* \left[h_1(x,\, a^*,\, s) + \beta\ \mathrm{E}\left[\omega(\ell(x,\, a^*, s),\, \varphi(\mu,\, \gamma^*),\, s')\right| s]\right] \\
= \;& \mu h_0(x,\, a^*,\, s) + \beta\ \mathrm{E}\left[\sum_{j=0}^{k} \mu^j \omega_j(\ell(x,\, a^*, s),\, \varphi(\mu,\, \gamma^*),\, s')\right| s\Bigg] \\
\geq \;& \mu h_0(x,\, \widetilde{a},\, s) + \beta\ \mathrm{E}\left[\sum_{j=0}^{k} \mu^j \omega_j(\ell(x,\, \widetilde{a},\, s),\, \varphi(\mu,\, \gamma^*),\, s')\right| s\Bigg] \\
+ \;& \gamma^* \left[h_1(x,\, \widetilde{a},\, s) + \beta\ \mathrm{E}\left[\omega(\ell(x,\, \widetilde{a}, s),\, \varphi(\mu,\, \gamma^*),\, s')\right| s]\right]
\end{aligned}$$

By assumption,

$$(\mu/\left\|\mu\right\|)h_0(x,\, a^*,\, s) + \beta\ \mathrm{E}\left[\sum_{j=0}^{k} (\mu^j/\left\|\mu\right\|)\omega_j(\ell(x,\, a^*, s),\, \varphi((\mu/\left\|\mu\right\|),\, (\gamma^*/\left\|\mu\right\|)),\, s')\right| s\Bigg]$$

is uniformly bounded (**A4** and $\omega \in M_b$ imply that there is uniform bound for the max value), while if $(\gamma^{*j}/\left\|\mu\right\|) > 0$ then

$$(\gamma^{*j}/\left\|\mu\right\|)\left[h_1^j(x,\, \widetilde{a},\, s) + \beta \mathrm{E}\left[\omega_j(\ell(x,\, \widetilde{a}, s),\, \varphi((\mu/\left\|\mu\right\|),\, (\gamma^*/\left\|\mu\right\|)),\, s')\right| s]\right] > 0;$$

therefore, there must be a $B > 0$ such that $\left\|\gamma^*\right\| \leq B \left\|\mu\right\|$ ■

Proposition 2. Let $\omega \in M_{bc}$ and assume **A1-A6** and **A7b.** There exists (a^*, γ^*) that solves $\mathbf{DSPP}_{(x,\mu,s)}$. Furthermore if **A6s** is assumed, then $a^*(x, \mu, s)$ is uniquely determined.

Proof: It is a relatively standard proof of existence of an equilibrium, based on a fixed point argument; see Appendix D.

The following Corollary to Theorem 3, is a simple restatement of the theorem in terms of the *Dynamic Saddle Point Problem*

Corollary 3.3. **($\mathbf{SPP}_\mu(x, s) \Longrightarrow \mathbf{DSPP}_{(x,\mu,s)}$).** Assume that $\mathbf{SPP}_\mu$ at (x, s) has a solution $(\boldsymbol{a}^*, \gamma^*)$ with value $V_\mu(x, s) = \sum_{j=0}^{l} \mu^j \omega_j^*(x, \mu, s)$, then (a_0^*, γ^*) solves $\mathbf{DSPP}_{(x,\mu,s)}$.

When $\mathbf{DSPP}_{(x,\mu,s)}$ has a solution, it defines a **SPFE** operator $T^* : \mathcal{M} \longrightarrow \mathcal{M}$ given by

$$(T^*W)(x,\mu,s) = \min_{\gamma\geq 0}\max_{a}\{\mu h_0(x,\,a,\,s) + \gamma h_1(x,\,a,\,s) + \beta\,\mathrm{E}\left[W(x',\,\mu',\,s')|\,s\right]\}$$
s.t. $x' = \ell(x,\,a,\,s),\ p(x,a,\,s) \geq 0$
and $\mu' = \varphi(\mu,\,\gamma)$,

When $W \equiv \mu\omega$, with $\omega \in M$, and $\mathbf{DSPP}_{(x,\mu,s)}$ uniquely defines the values $h_0^j(x,\,a^*(x,\,\mu,\,s),\,s)$, $j = 0,...,l$, then T^* defines a mapping, $T : M \longrightarrow M$, given by

$$(T\omega_j)(x,\,\mu,\,s) = h_0^j(x,\,a^*(x,\,\mu,\,s),\,s)+\beta\,\mathrm{E}\left[\omega_j(x^{*\prime}(x,\,\mu,\,s),\,\mu^{*\prime}(x,\,\mu,\,s),\,s')|\,s\right], \tag{40}$$

if $j = 0,...,k$, and

$$(T\omega_j)(x,\,\mu,\,s) = h_0^j(x,\,a^*(x,\,\mu,\,s),\,s),\text{ if } j = k+1,...,l.$$

Two remarks are in order. First, as already said, notice that $(T\omega_j)$ correspond to the 'promise keeping constraints' of the 'APG approach' to solving contractual problems but in our approach $(T\omega_j)$ is not a constraint: it is an outcome. Second, a fixed point of T^* does not imply a fixed point of T when 'the planner' is indifferent to T reallocations (e.g. $\mu_i = \mu_j$, $\mu_i(T\omega_i)+\ \mu_j(T\omega_j) = \text{constant}$) resulting in multiple (indeterminate) continuation values (for i and j)[18]

Proposition 3. Assume **DSPP** has a solution, for any $\omega \in M$ and (x,μ,s), then $T^* : \mathcal{M} \to \mathcal{M}$ is a well defined contraction mapping. Let $W^* = T^*(W^*)$ and $W^* = \mu\omega^*$, if in addition the solutions $a^*(x,\,\mu,\,s)$ to **DSPP** are unique, then $\omega^* = T(\omega^*)$ is unique.

Proof: The first part follows from showing that Blackwell's sufficiency conditions for a contraction are satisfied for T^* (see Lemmas 7A to 10A in Appendix D), the second part from the definition of T.

Our last Theorem, Theorem 5, wraps up our sufficiency results and is, in fact, a Corollary to Theorem 4. It shows how, starting from a *Dynamic Saddle-Point Problem* and a corresponding well defined *Contraction Mapping* resulting in a unique value function, one obtains the solution to our original problem $\mathbf{PP}_\mu$. The previous Propositions 2 and 3 provide conditions guaranteeing that the assumptions of Theorem 5, regarding T, are satisfied.

Theorem 5 ($\mathbf{DSPP}_{(x,\mu,s)} \Longrightarrow \mathbf{SPP}_\mu(x,s)$). Assume $T : M \to M$ has a unique fixed point ω^*. Then the *value function* $W^*(\mu,x,s) = \mu\omega^*(x,\mu,s)$ is the value of $\mathbf{SPP}_\mu$ at $(x,\,s)$ and the solutions of **DSPP** define a *saddle-point correspondence*Ψ, such that if $(\mathbf{a}^*,\boldsymbol{\gamma}^*)$ is generated by Ψ from $(x,\,\mu,\,s)$, then $(\mathbf{a}^*,\boldsymbol{\gamma}^*)$ solves $\mathbf{SPP}_\mu$ at $(x,\,s)$ and $\mathbf{a}^*$ is the unique solution to $\mathbf{PP}_\mu$ at $(x,\,s)$.

[18] Messner and Pavoni's 'counterxample' is one of these cases of 'flats in the Pareto frontier', when one only considers the T^* map and not the T map; see Marimon, Messner and Pavoni (2010).

Proof: By assumption **SPFE** is satisfied. The proof of Theorem 4 is based on having a unique representation $W^*(\mu, x, s) = \mu\omega^*(x, \mu, s)$ which, in that proof, is given by Lemma 2. In Theorem 5 such a unique representation is assumed, which implicitly also means to assume that the values $h_0^j(x,\, a^*(x,\, \mu,\, s),\, s)\, j = 0, ..., l$ are uniquely determined, which in fact is all what is needed in the proof of Theorem 4.

6 Related work

Precedents of our approach can be found in Kydland and Prescott (1980) and in Hansen, *et al.* (1985). They introduced lagrange multipliers as co-state variables. Our work provides a formal proof that a stationary optimal policy can be obtained by properly introducing lagrange multipliers as co-states. There are special cases, however, where there is a one-to-one relationship between states and co-states. Obviously, in these cases it is possible to obtain a policy function that only depends on the state variables, although it may be discontinuous[19]. Rustichini (1998) provides a recursive characterization that encompasses these cases. He focuses on deterministic models with one (default) constraint. In this class of models, in general, it is possible to reduce the dimensionality of the policy to its natural states[20]. Cooley, Quadrini and Marimon (2000) obtain a similar result in a stochastic model with possible default and risk neutral agents. Efficient contracts in these models have the special feature that if a state is reached there can no be different past promises, regarding the value of the contract at that state, depending on past contractual histories. In deterministic models past histories may be uniquely determined by the state and in models with risk neutral agents there is no need for consumption smoothing. However, in economies with uncertainty and risk-averse agents the effect of intertemporal (default) constraints must be *smoothed* and, as a result, since the same state may be reached following different contractual histories the optimal contract can not have a unique value associated with such state[21].

The pioneer work of Abreu, *et al.* (1990) -APS, from now on,- characterizing sub-game perfect equilibria, shows that past histories can be summarized in terms of promised utilities (we summarize them with the co-state μ). While the pioneer work of Green (1987) and Thomas and Worral (1988)) -GTW, from now on- shows that efficient contracts, promising a given initial level of (present value) utility, can have a recursive structure. These related approaches have

[19] In particular, discontinuities may arise at points where, given a value of the state variable, a co-state variable jumps. In contrast, the corresponding policy function of our approach (i.e., a function of states and co-states) may well be continuous. An example of such discontinuities can be found in Benhabib and Rustichini (1996).

[20] Notice that a similar "reduction" is proposed by the Subspace Method to solve, for example, Optimal Linear Regulator problems (see, for example, Hansen and Sargent (2000)). However, as the Subspace Method shows, even when the "reduction" is possible, can be very convenient to express policies in terms of state and co-state variables.

[21] See Marcet and Marimon (1992) or Kocherlakota (1995) for an explanation of how, to achieve consumption smoothing, idiosyncratic shocks imply increased consumption over the whole future. This can only be achieved by keeping track of past shocks through μ.

been widely used in macroeconomics[22]. In particular, Kocherlakota (1995) has applied GTW to characterize optimal social insurance contracts with participation constraints[23] and the APS approach has been further developed by Cronshaw and Luenberger (1994), to study dynamic games, and by Chang (1998) and Phelan and Stacchetti (1999), to study Ramsey models.

There are similarities and differences between our approach and the "promised utilities" approaches. The extent that these approaches duplicate, complement or dominate, each other will not be fully understood until they are more developed and applied[24]. Nevertheless, some conceptual differences emerge. For instance, our approach provides a common framework that encompasses two-period (dynamic competitive) constraints as well as discounted sums that may or may not be discounted utilities (as is the case with some present value budget constraints). Within this general framework, our approach directly characterizes efficient contracts, without having to find the whole set of feasible -incentive compatible- contracts. Presumably one could develop APS or GTW to provide a similar general framework (and the work of Chang (1998) and Phelan and Stacchetti (1999) are important steps in this direction), but it may not be the most efficient way to proceed. There are issues of dimensionality and *proper recursivity* that must be taken into account.

By *proper recursivity* we mean that our SPFE provides a recursive solution that starts out from pre-set initial conditions[25] In the GTW approach initial promised (present value) utilities must be specified for all -but one- agents and, therefore, needs to be known that it is feasible. As it is well known from standard Pareto Optimal problems such approach is very useful when there are two agents and the Pareto frontier is downward slopping, otherwise it can become very cumbersome[26]. Similarly, while the APS approach provides a method to obtain the set of feasible initial present values[27] in many applications may be a fairly roundabout way to proceed. For instance, one could apply the same method to solve a simple dynamic model -such as the neoclassical growth model-where interetemporal constraints are given by Euler equations. That is, given a

[22] Ljungqvist and Sargent (2000) provides an excellent introduction, and reference, to most of this recent work.

[23] A two agents (without capital) version of Marcet and Marimon (1992), recently studied using our approach by Attanasio and Rios-Rull (2000).

[24] For instance, our approach needs to be fully developed to incorporate private information constraints and its range of applications is yet to be exploited. Similarly, some of our assumptions -such as the linear additivity of utilities and constraints- need to be relaxed in order to make its generality at par with some of the developments of Rustichini, APS or GTW approaches. We leave these issues for future research (see, however, Footnote 2).

[25] For example, when we transform the problem **PP** into the problem **SPP** we set $\overline{\mu} = (1, 0, ..., 0)$.

[26] Even with two agents and a downward slopping Pareto frontier, as in Kocherlakota (1996), one may be interested in finding the efficient allocation that *ex-ante* gives the same utility to both agents. While this is trivial with our approach (just give the same initial weights in the **PP** problem), becomes very tricky with the GTW approach since the "right promise" must be made to determine the initial conditions.

[27] We refer to the *self-generation* and *factorization* properties of the operator characterizing incentive constraints (see, Abreu *et al.* (1990), and Chang (1998) and Ljungqvist and Sargent (2000) for macroeconomic applications).

"promised expected marginal utility" for next period, the Euler equation determines which current actions and marginal utilities are feasible. Proper iteration of such a map determines which initial conditions result in paths satisfying the *transversality conditions* and, therefore, the initial marginal utility that an efficient planner must set (given an initial capital and shock). Standard recursive methods, aimed at obtaining directly the value and policy functions, have proved to be a more useful approach for problems of this type. We think that the same principle applies to models with intertemporal constraints.

The strength of the APS approach, however, is in the study of models where one must know the set of feasible payoffs in order to characterize the efficient ones. In particular, as the work of Abreu (1988) shows, in order to characterize an efficient sub-game perfect equilibrium it is enough, in general, to know the *extremal points* of the set of feasible present value payoffs. More precisely, the *worst* equilibrium payoff may be sufficient to characterize the *best* equilibrium payoff, and a corresponding equilibrium path. Nevertheless, even in this class of models, it is often the case that one can separately obtain the worst sub-game perfect equilibrium payoff. If this is the case, such value plays the role of an intertemporal participation constraint in our framework. But, even when the computation of the *worst* equilibrium is not straightforward, our recursive approach, as a general method to obtain *extremal points,* may be a convenient approach[28]. In contrast, from a computational point of view, following the APS approach may not be very convenient general approach. In particular, while finding the set of feasible equilibrium payoffs may not be a hard computational problem when there is only one promised utility to consider, when there are n co-state variables a *set* in R^n has to be computed, which is not a straightforward computational problem. As we said, our approach is based on well specified initial values and deals with the computation of functions, which is a better understood problem[29].

Finally, an interesting -but not exclusive- feature of out approach is that the evolution of μ often helps to directly characterize the behavior of the model. For example, in models with participation constraints the μ's allow to interpret the behavior of the model as changing the Pareto weights sequentially depending on how binding the participation constraints become. In Ramsey type models the behavior of the μ's is associated with the commitment technology and the extent that markets are complete. For example, with full commitment and complete markets the μ's are constant after period one[30] Alternatively, in Aiyagari *et al.*

[28] *For example, Ljungqvist and Sargent (2000) show how separately compute the worst equilibrium in dynamic models with only one promised utility and no additional state variables. That our approach can also be used to obtain other extremal values may be seen by properly replacing the objective function, etc. Nevertheless we plan to make such application more explicit in future work.*

[29] For example, consider a version of the model of Chang (1998) with two agents. Certainly, an additional state variable needs to be introduced. If we use APS, the set of initial values would be a subset of R^2, and it would normally not be an interval, so it is difficult to characterize numerically. But using SPFE we know to set initial values for co-state variables to zero.

[30] For these models, our approach provides a recursive, alternative, formulation to the primal

(2002) the behavior of μ characterizes optimal tax policies as a risk-adjusted martingale. In summary, having optimal policies dependent of an additional co-state vector μ is not just convenient from a technical point of view, but also provides a fairly transparent interpretation of how intertemporal incentive constraints affect (efficient) economic outcomes.

7 Concluding remarks

We have shown that a large class of problems with implementability constraints can be analyzed by an equivalent recursive saddle point problem. This saddle point problem obeys a saddle point functional equation, which is a version of the Bellman equation. This approach works for a very large class of models with incentive constraints, restricted budget constraint, optimal policy, optimal regulation, etc. This means that a unified framework can be provided to analyze all these models. The key feature of our approach is that instead of having to write optimal contracts as history-dependent contracts one can write them as a stationary function of the standard state variables together with additional co-state variables. These co-state variables are -recursively- obtained from the Lagrange multipliers associated with the intertemporal incentive constraints, starting from pre-specified initial conditions. With such approach we aim to extends the existing set of tools available to study dynamic economies with intertemporal constraints.

Our current research aims at relaxing several of the assumptions, in particular that of full information, developing in detail some computational aspects of this method, and exploring a range of applications to several models, including strategic dynamic behavior, optimal policy and borrowing under incomplete insurance.

approach developed in Lucas and Stokey (1983) and Chari, *et al.* (1995).

8 Authors' Affiliations

Albert Marcet, London School of Economics and Institut d'Analisi Economica,
E-mail: a.marcet@lse.ac.uk
and
Ramon Marimon, European University Institute and Universitat Pompeu Fabra - BarcelonaGSE,
Via delle Fontanelle 20, I-50014 San Domenico di Fiesole, Italy;
Telephone: +39-055-4685809, Fax: +39-055-4685894,
E-mail: ramon.marimon@eui.eu

APPENDIX

Appendix A (Proof of Proposition 1)

The proof of Proposition 1 relies on the following result:

Lemma 1A. Assume **A1-A6** and that S is countable**,** then

i) $\mathcal{B}(x,s)$ is non-empty, convex, bounded and $\sigma(\mathcal{L}_\infty,\, \mathcal{L}_1)$ closed; therefore it is $\sigma(\mathcal{L}_\infty,\, \mathcal{L}_1)$ compact

ii) Given $d \in R$, the set $\left\{\boldsymbol{a} \in \mathcal{A} : f_{(x,\mu,s)}(\boldsymbol{a}) \geq d\right\}$ is convex and $\sigma(\mathcal{L}_\infty,\, \mathcal{L}_1)$ closed

The proof of Lemma 1A builds on three theorems. First, the *Urysohn metrization theorem* stating that regular topological spaces with a countable base are metrizable[31]. Second, the *Mackey-Arens theorem* stating that different topologies consistent with the same duality share the same closed convex sets; in our case, the duality is $(\mathcal{L}_\infty,\, \mathcal{L}_1)$ and the weakest and the strongest topology consistent with such duality; namely, the *weak-star*, $\sigma(\mathcal{L}_\infty,\, \mathcal{L}_1)$ and the Mackey $\tau(\mathcal{L}_\infty,\, \mathcal{L}_1)$. Third, the *Alaoglu theorem* stating that norm bounded $\sigma(\mathcal{L}_\infty,\, \mathcal{L}_1)$ closed subsets of $\mathcal{L}_\infty$ are $\sigma(\mathcal{L}_\infty,\, \mathcal{L}_1)$ compact[32].

Proof:

Assumptions **A2**, and **A4 - A6** imply that $\mathcal{B}(x,s)$ is convex, and closed under pointwise convergence. Since, by assumption S is countable, *Urysohn metrization theorem* guarantees that $\mathcal{B}(x,s)$ is, in fact, $\sigma(\mathcal{L}_\infty,\, \mathcal{L}_1)$ closed. Assumptions **A3** and **A4** imply that $\mathcal{B}(x,s)$ is bounded in the $\|\cdot\|_\infty^\beta$ norm as needed for compactness, according to the *Alaoglu theorem.*

Assumptions **A5** and **A6** imply that $\mathcal{B}(x,s)$ and the upper contour sets $\{\boldsymbol{a} \in \mathcal{A} : f_\mu(\boldsymbol{a}) \geq d\}$, are convex and, together with the previous assumptions (**A2**, and **A4**), Mackey closed and, therefore, $\sigma(\mathcal{L}_\infty,\, \mathcal{L}_1)$ closed[33]■

Proof of Proposition 1: As in Bewley (1991), the central element of the proof follows from the *Hausdorff maximal principle* and an application of *the finite intersection property*[34]. Let $\mathcal{P}_d = \left\{\boldsymbol{a} \in \mathcal{B}(x,s) : f_{(x,\mu.s)}(\boldsymbol{a}) \geq d\right\}$,

[31]See Dunford and Schwartz (1957) p. 24. in our case the metric we use is given by

$$\rho_\infty^\beta(\boldsymbol{a},\boldsymbol{b}) = \sum_{n=0}^{\infty} \beta^n \sup_{s^n \in S^n} \mid a_n(s) - b_n(s) \mid .$$

[32]See Schaefer (1966) p. 130 and p. 84, respectively.

[33]See Bewley (1972) for a proof of the Mackey continuity expected utility, without assuming S being countable.

[34]See, Kelley (1955) p. 33-34. for the Hausdorff principle and the Minimal principle, and p. 136 for the theorem stating that "a set is compact if and only if every family of closed sets which has a the finite intersection property has a non-void intersection."

then by Lemma 1A , $\mathcal{P}_d$ is $\sigma(\mathcal{L}_\infty, \mathcal{L}_1)$ closed. By the interiority assumption of Proposition 1, for d low enough, it is non-empty. In fact, we can consider the family of sets $\{\mathcal{P}_d : d \in D\}$ for which $\mathcal{P}_d \neq \varnothing$, where $D \subset R$. The sets $\mathcal{P}_d$ are ordered by inclusion; in fact, if $d' > d$ then $\mathcal{P}_{d'} \subset \mathcal{P}_d$ and every finite collection of them has a non-empty intersection (i.e., $\{\mathcal{P}_d : d \in D\}$ satisfies *the finite intersection property*), but then by compactness of $\mathcal{B}(x,s)$ any family of subsets of $\{\mathcal{P}_d : d \in D\}$ -say, $\{\mathcal{P}_d : d \in B \subset D\}$- has a non-empty intersection and, by inclusion, there is $\mathcal{P}_{\widehat{d}} = \cap \{\mathcal{P}_d : d \in B \subset D\} \neq \varnothing$. In particular, there is $\mathcal{P}_{d^*} = \cap \{\mathcal{P}_d : d \in D\} \neq \varnothing$ which -as the *the minimal principle* states- is a minimal member of the family $\{\mathcal{P}_d : d \in D\}$. It follows that if $\boldsymbol{a}^* \in \mathcal{P}_{d^*}$ then $f_{(x,\mu.s)}(\boldsymbol{a}^*) \geq f_{(x,\mu.s)}(\boldsymbol{a})$ for any $\boldsymbol{a} \in \mathcal{B}(x,s)$. Furthermore, if strictly concavity is assumed then $\mathcal{P}_{d^*}$ must be a singleton, otherwise convex combinations of elements of $\mathcal{P}_{d^*}$ will form a proper closed subset of $\mathcal{P}_{d^*}$ contradicting its minimality■

Appendix B (Some Properties of W^*)

Lemma 2A. Let $W^*(x,\mu,s) \equiv V_\mu(x,s)$ be the value of $\mathbf{SPP}_\mu$ at (x,s), for an arbitrary $(x,\ \mu,\ s)$, then

i) $W^*(x,\cdot,s)$ is convex and homogeneous of degree one;

ii) if **A2**- **A4** are satisfied $W^*(\cdot,\mu,s)$ is continuous and uniformly bounded, and

iii) if **A5** and **A6** are satisfied $W^*(\cdot,\mu,s)$ is concave.

Proof: *i*) follows from the fact that, for any $\lambda > 0$, $f_{(x,\lambda\mu.s)}(\boldsymbol{a}) = \lambda f_{(x,\mu.s)}(\boldsymbol{a})$. To see this, let $(\gamma^*, \boldsymbol{a}^*)$ satisfy **SFPE**, i.e.

$$\begin{aligned} & f_{(x,\mu.s)}(\boldsymbol{a}^*) + \gamma g(\boldsymbol{a}^*)_0 \\ \geq \quad & f_{(x,\mu.s)}(\boldsymbol{a}^*) + \gamma^* g(\boldsymbol{a}^*)_0 \\ \geq \quad & f_{(x,\mu.s)}(\boldsymbol{a}) + \gamma^* g(\boldsymbol{a})_0, \end{aligned}$$

for any $\gamma \in R_+^{l+1}$ and $\boldsymbol{a} \in \mathcal{B}'(x,s)$. Then $(\lambda\gamma^*, \boldsymbol{a}^*)$ satisfies

$$\begin{aligned} & f_{(x,\lambda\mu.s)}(\boldsymbol{a}^*) + \gamma g(\boldsymbol{a}^*)_0 \\ \geq \quad & f_{(x,\lambda\mu.s)}(\boldsymbol{a}^*) + \lambda\gamma^* g(\boldsymbol{a}^*)_0 \\ \geq \quad & f_{(x,\lambda\mu.s)}(\boldsymbol{a}) + \lambda\gamma^* g(\boldsymbol{a})_0, \end{aligned}$$

for any $\gamma \in R_+^{l+1}$ and $\boldsymbol{a} \in \mathcal{B}'(x,s)$. *ii*) and *iii*) are straightforward; in particular, *ii*) follows from applying the Theorem of the Maximum (Stokey, Lucas and Prescott,1989, Theorem 3.6) *iii*) follows from the fact that the constraint sets are convex and the objective function concave.■

Lemma 3A: If the inf sup problem, **SPFE** at $(x,\ \mu,\ s)$, has a solution then the value of this solution is unique.

Proof: It is a standard argument: consider two solutions to the right side of **SPFE** at (x, μ, s), $(\widetilde{a}, \widetilde{\gamma})$ and $(\widehat{a}, \widehat{\gamma})$, then repeated application of the saddle-point condition implies:

$$\begin{aligned} & \mu h_0(x, \widetilde{a}, s) + \widetilde{\gamma} h_1(x, \widetilde{a}, s) + \beta \mathrm{E}\left[W^*(\ell(x, \widetilde{a}, s'), \varphi(\mu, \widetilde{\gamma}), s') | s\right] \\ \geq\ & \mu h_0(x, \widehat{a}, s) + \widetilde{\gamma} h_1(x, \widehat{a}, s) + \beta \mathrm{E}\left[W^*(\ell(x, \widehat{a}, s'), \varphi(\mu, \widetilde{\gamma}), s') | s\right] \\ \geq\ & \mu h_0(x, \widehat{a}, s) + \widehat{\gamma} h_1(x, \widehat{a}, s) + \beta \mathrm{E}\left[W^*(\ell(x, \widehat{a}, s'), \varphi(\mu, \widehat{\gamma}), s') | s\right] \\ \geq\ & \mu h_0(x, \widetilde{a}, s) + \widehat{\gamma} h_1(x, \widetilde{a}, s) + \beta \mathrm{E}\left[W^*(\ell(x, \widetilde{a}, s'), \varphi(\mu, \widehat{\gamma}), s') | s\right] \\ \geq\ & \mu h_0(x, \widetilde{a}, s) + \widetilde{\gamma} h_1(x, \widetilde{a}, s) + \beta \mathrm{E}\left[W^*(\ell(x, \widetilde{a}, s'), \varphi(\mu, \widetilde{\gamma}), s') | s\right] \end{aligned}$$

therefore the value of the objective at both $(\widetilde{a}, \widetilde{\gamma})$ and $(\widehat{a}, \widehat{\gamma})$ coincides ■

Appendix C (Some Properties of convex homogeneous functions)

To simplify the exposition of these properties let $F : R^m_+ \rightarrow R$ continuous, convex and homogeneous of degree one. The *subgradient set* of F at y, denoted $\partial F(y)$, is given by

$$\partial F(y) = \left\{ z \in R^m \mid F(y') \geq F(y) + (y' - y)z \text{ for all } y' \in R^m_+ \right\}.$$

The following **facts**, regarding F, are used in proving Lemmas 1 and 2:

F1. If F is convex, then it is differentiable at y if, and only if, $\partial F(y)$ consists of a single vector; i.e. $\partial F(y) = \{\nabla F(y)\}$, where $\nabla F(y)$ is called the *gradient* of F at y.

F2. If F is convex and finite in a neighborhood of y, then $\partial F(y)$ is the convex hull of the compact set

$$\{z \in R^m \mid \exists\, y_k \longrightarrow y \text{ with } F \text{ differentiable at } y_k \text{ and } \nabla F(y_k) \longrightarrow z\}.$$

F3. Lemma 4A (Euler's formula). If F is convex and homogeneous of degree one and $z \in \partial F(y)$ then $F(y) = yz$.

F4. Lemma 5A. If F is convex and homogeneous of degree one, for any pair $\left(f, \widehat{f}\right)$, if $f^d(y) \in \partial F(y)$ and $f^d(\widehat{y}) \in \partial F(\widehat{y})$, then $\widehat{y} f^d(\widehat{y}) \geq \widehat{y} f^d(y)$.

F1 is a basic result on differentiability of convex functions (see, Rockafellar, 1981, Theorem 4F, or 1970, Theorem 25.1). **F2** is a very convenient characterization of the subgradient set of a convex function (see Rockafellar, 1981, Theorem 4D, or 1970, Theorem 25.6). We now provide a proof of the last two facts.

Proof of Lemma 4A: Let $z \in \partial F(y)$, then for any $\lambda > 0$, $F(\lambda y) - F(y) \geq (\lambda y - y)z$, and, by homogeneity of degree one: $(\lambda - 1)F(y) \geq (\lambda - 1)yz$. If $\lambda > 1$ this weak inequality results in: $F(y) \geq yz$; while if $\lambda \in (0, 1)$ in $F(y) \leq yz$■[35]

[35] Notice that this does not imply that F is linear, which requires that $F(-y) = -F(y)$.

Proof of Lemma 5A: To see **F4** notice that if $f^d(y) \in \partial F(y)$ and $f^d(\widehat{y}) \in \partial F(\widehat{y})$, by convexity, homogeneity of degree one, and Euler's formula: $F(\widehat{y}) = \widehat{y} f^d(\widehat{y})$ and

$$\begin{aligned} F(\widehat{y}) &\geq F(y) + (\widehat{y} - y) f^d(y) \\ &= y f^d(y) + \widehat{y} f^d(y) - y f^d(y) = \widehat{y} f^d(y). \end{aligned}$$

Appendix D (Proof of Propositions 2 and 3)

Proof of Proposition 2: Given the assumptions of Proposition 2, for any (x, μ, s), and $\gamma \in \mathcal{R}_+^{l+1}$, let

$$F_{(x,\mu,s)}(\gamma) = \arg\sup_a \left\{ \mu h_0(x,\, a,\, s) + \beta \ \mathrm{E}\left[\sum_{j=0}^{k} \mu^j \omega_j(x',\, \mu',\, s') |\, s \right] \right\}$$

s.t. $x' = \ell(x,\, a,\, s),\ p(x, a,\, s) \geq 0$

and $\mu' = \varphi(\mu,\, \gamma)$,

$$\text{and } h_1^j(x, a, s) + \beta \mathrm{E}\left[\omega^j(x', \mu', s') |\, s \right] \geq 0,\ j = 0, ..., l \tag{41}$$

Since this is a standard maximization problem of a continuous function on a compact set, there is a solution $a^*(x, \mu, s; \gamma) \in F_{(x,\mu,s)}(\gamma)$. Furthermore, given that the constraint set is convex and has a non-empty interior (by **A2** and **A7b**), there is an associated multiplier vector; let $\gamma^{*j}(x, \mu, s; \gamma)$ be the multiplier corresponding to (**??**), for j. In particular, by Lemma 3, $\gamma^*(x, \mu, s; \gamma) \in G_{(x,\mu,s)}(a^*)$, where

$$\begin{aligned} G_{(x,\mu,s)}(a) &= \arg \inf_{\{\gamma \geq 0: \|\gamma\| \leq B\|\mu\|\}} \{ \mu h_0(x,\, a,\, s) + \gamma h_1(x,\, a,\, s) + \beta \ \mathrm{E}\left[\varphi(\mu,\, \gamma) \omega(x',\, \varphi(\mu,\, \gamma),\, s') |\, s \right] \} \\ \text{s.t. } x' &= \ell(x,\, a,\, s),\ p(x, a,\, s) \geq 0 \end{aligned}$$

The rest of the proof is a trivial application of the *Theorem of the Maximum* (e.g. Stokey et al. (1989), p. 62) and of *Kakutani's Fixed Point Theorem* (e.g. Mas-Colell et al. (1995), p.953). First notice that $\mu h_0(x,\, a,\, s) + \beta \ \mathrm{E}\left[\sum_{j=0}^{k} \mu^j \omega_j(x',\, \mu',\, s') |\, s \right]$ is continuous in a and, by **A2**, **A4** and the definition of M_{bc}, $G^*_{(x,\mu,s)}(\cdot)$ is also coninuous. Second, let $A(x, s) \equiv \{a \in A : p(x, a,\, s) \geq 0\}$, by **A2** and **A3** $A(x; s)$ is compact and by **A5** is convex, while $\Psi(\mu) \equiv \{\gamma \geq 0 : \|\gamma\| \leq B \|\mu\|\}$ is trivially compact and convex. Therefore, $F_{(x,\mu,s)} : \Psi(\mu) \to A(x; s)$ and $G^*_{(x,\mu,s)} : A(x; s) \to \Psi(\mu)$ are upper-hemicontinuous, non-empty and convex-valued correspondences, jointly mapping a convex and compact set, $\Psi \times A(x; s)$, in itself; by *Kakutani's Fixed Point Theorem* there is a fixed point (a^*, γ^*) which is a solution to $\mathbf{DSPP}_{(x,\mu,s)}$. Furthermore, $F_{(x,\mu,s)}(\cdot)$ is a continuous function, when **A6s** is assumed■

Lemma 6A. $\mathcal{M}$ is a nonempty complete metric space.

Proof: That it is non-empty is trivial. Except for the homogeneity property, that every Cauchy sequence $\{W^n\} \in \mathcal{M}_{bc}$ converges to $W \in \mathcal{M}_{bc}$ satisfying i), iii), and the convexity property ii), follows from standard arguments (see, for example, Stokey, et al. (1989), Theorem 3.1 and Lemma 9.5). To see that the homogeneity property is also satisfied, for any $(x,\, \mu,\, s)$ and $\lambda > 0$,

$$\begin{aligned}
& |W(x,\, \lambda\mu,\, s) - \lambda W(x,\, \mu,\, s)| \\
= \;& |W(x,\, \lambda\mu,\, s) - W^n(x,\, \lambda\mu,\, s) + \lambda W^n(x,\, \mu,\, s) - \lambda W(x,\, \mu,\, s)| \\
\leq \;& |W(x,\, \lambda\mu,\, s) - W^n(x,\, \lambda\mu,\, s)| + \lambda |W^n(x,\, \mu,\, s) - W(x,\, \mu,\, s)| \\
\rightarrow \;& 0
\end{aligned}$$

■

Lemma 7A. Assume **A2 - A6** and **A7b**.The operator T^* maps $\mathcal{M}_{bc}$ into itself.

Proof: First, notice that by Proposition 2, given $W \in \mathcal{M}_{bc}$, T^*W is well defined. The correspondences $\Gamma : X \rightarrow X$ and $\Phi : \mathcal{R}_+^{l+1} \rightarrow \mathcal{R}_+^{l+1}$, defined by $\Gamma(x)_{(\mu,s)} \equiv \{x' \in X : x' = \ell(x,\, a,\, s),\ p(x, a, s) \geq 0, \text{ for some } a \in A\}$ and $\Phi(\mu)_{(x,s)} \equiv \left\{\mu' \in \mathcal{R}_+^{l+1} : \mu' = \varphi(\mu,\, \gamma), \text{ for some } \gamma \in \mathcal{R}_+^{l+1}\right\}$, are continuous and compact-valued (by **A2**, **A3** and **A5,** and the definition of φ, respectively) and, as in the proof of Proposition 2, by continuity of the objective function, it follows that $T^*W(\cdot,\, \cdot,\, s)$ is continuous, and given **A3** and **A4**, and the boundedness condition on W, it follows that T^*W also satisfies (i).To see that the homogeneity properties are satisfied, let $(a^*,\, \gamma^*)$ satisfy

$$(T^*W)(x,\, \mu,\, s) = \mu h_0(x,\, a^*,\, s) + \gamma^* h_1(x,\, a^*,\, s) + \beta \mathrm{E} W(x^*,\, \mu^*,\, s')$$

then, for any $\lambda > 0$

$$\lambda(T^*W)(x,\, \mu,\, s) = \lambda[\mu h_0(x,\, a^*,\, s) + \gamma^* h_1(x,\, a^*,\, s) + \beta \mathrm{E} W(x^*,\, \mu^*,\, s')]$$

Furthermore,

$$\begin{aligned}
& \lambda\mu h_0(x,\, a^*,\, s) + \lambda\gamma^* h_1(x,\, a^*,\, s) + \beta \mathrm{E} W(x^*,\, \lambda\mu^{*'},\, s') \\
= \;& \lambda \left[\mu h_0(x,\, a^*,\, s) + \gamma^* h_1(x,\, a^*,\, s) + \beta \mathrm{E} W(x^*,\, \mu^{*'},\, s')\right]
\end{aligned}$$

To see that **SPFE** is satisfied, let $\gamma \geq 0$, $\mu' = \varphi(\lambda\mu,\, \gamma)$ and $a \in A(x, s)$, $x' = \ell(x,\, a,\, s')$, then

$$\begin{aligned}
& \lambda\mu h_0(x,\, a^*,\, s) + \gamma h_1(x,\, a^*,\, s) + \beta \mathrm{E} W(x^*,\, \mu',\, s') \\
= \;& \lambda \left[\mu h_0(x,\, a^*,\, s) + \gamma\lambda^{-1} h_1(x,\, a^*,\, s) + \beta \mathrm{E} W(x^*,\, \mu'\lambda^{-1},\, s')\right] \\
\geq \;& \lambda \left[\mu h_0(x,\, a^*,\, s) + \gamma^* h_1(x,\, a^*,\, s) + \beta \mathrm{E} W(x^*,\, \mu^{*'},\, s')\right] \\
\geq \;& \lambda \left[\mu h_0(x,\, a,\, s) + \gamma^* h_1(x,\, a,\, s) + \beta \mathrm{E} W(x',\, \mu^{*'},\, s')\right]
\end{aligned}$$

It follows that,

$$\begin{aligned}(T^*W)(x,\,\lambda\mu,\,s) &= \lambda\mu h_0(x,\,a^*,\,s)+\lambda\gamma^* h_1(x,\,a^*,\,s)+\,\beta \mathrm{E}W(x^*,\;\lambda{\mu^*}',\,s')\\ &= \lambda(T^*W)(x,\,\mu,\,s)\end{aligned}$$

Finally, since $W\in\mathcal{M}_{bc}$, it is straightfoward to show that TW is concave in x (by **A5** and **A6),** and convex in μ■

Lemma 7A (monotonicity) Let $\widehat{W}\in\mathcal{M}$ and $\widetilde{W}\in\mathcal{M}$ be such that $\widehat{W}\leq\widetilde{W}$, then $(T^*\widehat{W})\leq(T^*\widetilde{W})$.

Proof Fix $(\mu,\,x,\,s)$, then for any μ' satisfying $\mu'=\varphi(\mu,\,\gamma)\geq 0$, for a given $\gamma\geq 0$,

$$\begin{aligned}&\max_{a\in A(x,s)}\{\mu h_0(x,\,a,\,s)+\gamma h_1(x,\,a,\,s)+\beta\mathrm{E}\widehat{W}(\ell(x,\,a,\,s),\mu',\,s')\}\\ \leq\;&\max_{a\in A(x,s)}\{\mu h_0(x,\,a,\,s)+\gamma h_1(x,\,a,\,s)+\beta\mathrm{E}\widetilde{W}(\ell(x,\,a,\,s),\mu',\,s')\}\end{aligned}$$

It follows that

$$\begin{aligned}&\min_{\gamma\geq 0}\max_{a\in A(x,s)}\{\mu h_0(x,\,a,\,s)+\gamma h_1(x,\,a,\,s)+\beta\mathrm{E}\widehat{W}(\ell(x,\,a,\,s),\varphi(\mu,\,\gamma),\,s')\}\\ \leq\;&\min_{\gamma\geq 0}\max_{a\in A(x,s)}\{\mu h_0(x,\,a,\,s)+\gamma h_1(x,\,a,\,s)+\beta\mathrm{E}\widetilde{W}(\ell(x,\,a,\,s),\varphi(\mu,\,\gamma),\,s')\}\end{aligned}$$

■

Notice that if $W\in\mathcal{M}_{bc}$ and $r\in\mathcal{R}$, $(W+r)(x,\,\mu,\,s)=\mu(\omega+r)(x,\,\mu,\,s)=\mu\omega(x,\,\mu,\,s)\,+\,r\,\|\mu\|$.

Lemma 8A (discounting) Assume **A4** and **A7b**. For any $W\in\mathcal{M}_b$, and $r\in\mathcal{R}_+$, $T^*(W\,+\,r)\leq T^*W\,+\,\beta r$.

Proof First notice that, for any $(x,\,\mu,\,s)$ and $\gamma\geq 0$,

$$\begin{aligned}&\max_{a\in A(x,s)}\{\mu h_0(x,\,a,\,s)+\gamma h_1(x,\,a,\,s)+\beta\mathrm{E}(W+r)(\ell(x,\,a,\,s),\varphi(\mu,\,\gamma),\,s')\}\\ =\;&\max_{a\in A(x,s)}\{\mu h_0(x,\,a,\,s)+\gamma h_1(x,\,a,\,s)+\beta\mathrm{E}W(\ell(x,\,a,\,s),\varphi(\mu,\,\gamma),\,s')+\beta r\,\|\varphi(\mu,\,\gamma)\|\}\\ =\;&\max_{a\in A(x,s)}\{\mu h_0(x,\,a,\,s)+\gamma h_1(x,\,a,\,s)+\beta\mathrm{E}W(\ell(x,\,a,\,s),\varphi(\mu,\,\gamma),\,s')\}+\beta r\,\|\varphi(\mu,\,\gamma)\|\,,\end{aligned}$$

Now, given $(x,\,\mu,\,s)$ and $a\in A(x,s)$, denote by $\gamma^+(a)$ the solution to the following problem

$$\begin{aligned}&\min_{\{\gamma\geq 0:\|\gamma\|\leq B\|\mu\|\}}\left\{\begin{array}{c}\mu h_0(x,\,a,\,s)+\beta\mathrm{E}\sum_{j=0}^k\mu^j(\omega_j+r)(\ell(x,\,a,\,s),\varphi(\mu,\,\gamma),\,s')\\ +\sum_{j=0}^l\gamma^j\left[h_1^j(x,\,a,\,s)+\beta\mathrm{E}(\omega_j+r)(\ell(x,\,a,\,s),\varphi(\mu,\,\gamma),\,s')\right]\end{array}\right\}\\ =\;&\mu h_0(x,\,a,\,s)+\beta\mathrm{E}\sum_{j=0}^k\mu^j\omega_j(\ell(x,\,a,\,s),\varphi(\mu,\,\gamma^+(a)),\,s')\\ +\;&\gamma^+(a)\left[h_1(x,\,a,\,s)+\beta\mathrm{E}\omega(\ell(x,\,a,\,s),\varphi(\mu,\,\gamma^+(a)),\,s')\right]+\beta r\left\|\varphi(\mu,\,\gamma^+(a))\right\|\end{aligned}$$

and let $\gamma^*(a)$ be the solution to

$$
\begin{aligned}
& \min_{\{\gamma\geq 0:\|\gamma\|\leq B\|\mu\|\}} \left\{ \begin{array}{c} \mu h_0(x,\, a,\, s) + \beta \mathrm{E} \sum_{j=0}^{k} \mu^j \omega_j(\ell(x,\, a,\, s), \varphi(\mu,\, \gamma),\, s') \\ \sum_{j=0}^{l} \gamma^j \left[h_1^j(x,\, a,\, s) + \beta \mathrm{E} \omega_j(\ell(x,\, a,\, s), \varphi(\mu,\, \gamma),\, s') \right] \end{array} \right\} \\
= \quad & \mu h_0(x,\, a,\, s) + \beta \mathrm{E} \sum_{j=0}^{k} \mu^j \omega_j(\ell(x,\, a,\, s), \varphi(\mu,\, \gamma^*(a)),\, s') \\
+ \quad & \gamma^*(a) \left[h_1(x,\, a,\, s) + \beta \mathrm{E} \omega(\ell(x,\, a,\, s), \varphi(\mu,\, \gamma^*(a)),\, s') \right];
\end{aligned}
$$

therefore,

$$
\begin{aligned}
& \mu h_0(x,\, a,\, s) + \beta \mathrm{E} \sum_{j=0}^{k} \mu^j \omega_j(\ell(x,\, a,\, s), \varphi(\mu,\, \gamma^+(a)),\, s') \\
+ \quad & \gamma^+(a) \left[h_1(x,\, a,\, s) + \beta \mathrm{E} \omega(\ell(x,\, a,\, s), \varphi(\mu,\, \gamma^+(a)),\, s') \right] + \beta r \left\| \varphi(\mu,\, \gamma^+(a)) \right\| \\
\leq \quad & \mu h_0(x,\, a,\, s) + \beta \mathrm{E} \sum_{j=0}^{k} \mu^j \omega_j(\ell(x,\, a,\, s), \varphi(\mu,\, \gamma^*(a)),\, s') \\
+ \quad & \gamma^*(a) \left[h_1(x,\, a,\, s) + \beta \mathrm{E} \omega(\ell(x,\, a,\, s), \varphi(\mu,\, \gamma^*(a)),\, s') \right] + \beta r \left\| \varphi(\mu,\, \gamma^*(a)) \right\|.
\end{aligned}
$$

Piecing things together (denoting, as usual, by $a^* \equiv a^*(x, \mu,\, s)$ and $\gamma^* \equiv \gamma^*(x, \mu,\, s)$); that is, $\gamma^* \equiv \gamma^*(a^*)$), we have:

$$
\begin{aligned}
& T^*(W\, +\, r)(x, \mu,\, s) \\
= \quad & \min_{\gamma\geq 0} \max_{a\in A(x,s)} \left\{ \begin{array}{c} \mu h_0(x,\, a,\, s) + \beta \mathrm{E} \sum_{j=0}^{k} \mu^j \omega_j(\ell(x,\, a,\, s), \varphi(\mu,\, \gamma),\, s') \\ + \sum_{j=0}^{l} \gamma^j \left[h_1^j(x,\, a,\, s) + \beta \mathrm{E} \omega_j(\ell(x,\, a,\, s), \varphi(\mu,\, \gamma),\, s') \right] \end{array} \right\} \\
\leq \quad & \mu h_0(x,\, a^*,\, s) + \beta \mathrm{E} \sum_{j=0}^{k} \mu^j \omega_j(\ell(x,\, a^*,\, s), \varphi(\mu,\, \gamma^*,\, s') \\
+ \quad & \gamma^* \left[h_1(x,\, a^*,\, s) + \beta \mathrm{E} \omega(\ell(x,\, a^*,\, s), \varphi(\mu,\, \gamma^*),\, s') \right] + \beta r \left\| \varphi(\mu,\, \gamma^*) \right\| \\
= \quad & \mu h_0(x,\, a^*,\, s) + \beta \mathrm{E} \sum_{j=0}^{k} \mu^j \omega_j(\ell(x,\, a^*,\, s), \varphi(\mu,\, \gamma^*,\, s') + \beta r \left\| \varphi(\mu,\, \gamma^*) \right\| \\
\leq \quad & \mu h_0(x,\, a^*,\, s) + \beta \mathrm{E} \sum_{j=0}^{k} \mu^j \omega_j(\ell(x,\, a^*,\, s), \varphi(\mu,\, \gamma^*,\, s') + \beta r (1+B) \left\| \mu \right\|
\end{aligned}
$$

By homogeneity, without loss of generality we can choose an arbitrary $\mu \neq 0$, such that: $\|\mu\| \leq (1+B)^{-1}$. The above inequalities show that $T^*(W\, +\, r) \leq T^*(W) + \beta r$■

Lemma 9A (Contraction property): The argument is the standard Blackwell's argument. We show that the contraction property is satisfied. Let $W, \widehat{W} \in \mathcal{M}_{bc}$.

Notice that $W \leq \widehat{W} + \left\| W - \widehat{W} \right\|$, then using the results of Lemmas 7A and 8A,

$$T^{*}W \leq T^{*}(\widehat{W} + \left\| W - \widehat{W} \right\|) \leq T^{*}(\widehat{W}) + +\beta \left\| W - \widehat{W} \right\|),$$

reversing the roles of W and $\widehat{W}$, we obtain that

$$\left\| T^{*}W - T^{*}\widehat{W} \right\| \leq \beta \left\| W - \widehat{W} \right\|;$$

therefore, T^{*} is a contraction mapping$\blacksquare$

REFERENCES

Abreu, D., D. Pearce and E. Stacchetti (1990): "Toward a Theory of Discounted Repeated Games with Imperfect Monitoring," *Econometrica* 58, 1041–1063.

Acemoglu, D., M. Golosov and A. Tsyvinskii (2011): "Power Fluctuations and Political Economy ", *Journal of Economic Theory* (forthcoming)

Aiyagari, R., A. Marcet, T. J. Sargent and J. Seppala (2002): "Optimal Taxation without State-Contingent Debt", *Journal of Political Economy* 110, 1220-54.

Alvarez, F. and N. L. Stokey (1998): "Dynamic Programming with Homogeneous Functions." *Journal of Economic Theory* 82(1), 167-89.

Alvarez, F. and U.J. Jerman (1998): "Efficiency, Equilibrium and Asset Pricing with Risk of Default." *Econometrica* 68, 775-98. 82(1), 167-89.

Attanasio, O. and J.-V. Rios-Rull (2000): "Consumption Smoothing in Island Economies: Can Public Insurance Reduce Welfare?," *European Economic Review* 44, 1225-58.

Bewley, T. F. (1972): "Existence of Equilibria in Economies with Infinitely Many Commodities," *Journal of Economic Theory* 4: 514–540.

Bewley, T. F. (1991): "A Theorem on the Existence of Competitive Equilibria in a Market with a Finite Number of Agents and Whose Commodity Space is $\mathbf{L}_\infty$", in *Equilibrium Theory in Infinite Dimensional Spaces* ed. by M.A. Khan and N.C. Yannelis. Springer-Verlag.

Chang, R. (1998): "Credible Monetary Policy with Long-Lived Agents: Recursive Approaches," *Journal of Economic Theory* 81, 431-461.

Chari, V.V., Christiano, L.J. and Kehoe, P.J.(1995): "Policy Analysis in Business Cycle Models." *Frontiers of Business Cycle Research*, ed. by T. F. Cooley, Princeton, N.J.: Princeton University Press.

Cooley, T. F. (ed.) (1995): *Frontiers of Business Cycle Research.* Princeton, N.J.: Princeton University Press.

Cooley, T. F, R. Marimon and V. Quadrini (2004): "Optimal Financial Contracts with Limited Enforceability and the Business Cycle," *Journal of Political Economy* 112, 817-47.

Cronshaw, M. and D. Luenberger (1994): "Strongly Symmetric Subgame Perfect Equilibria in Infinitely Repeated Games with Perfect Monitoring and Discounting," *Games and Economic Behavior,* 6, 220–237.

Dunford, N. and J. T. Schwartz (1957): *Linear Operators, Part I.* New York: John Wiley & Sons.

Green, E. J. (1987): "Lending and the Smoothing of Uninsurable Income" in *Contractual Arrangements for Intertemporal Trade* ed. by E. C. Prescott and N. Wallace, University of Minnesota Press.

Kehoe, P. and F. Perry (2002): "International Business Cycles with Endogenous Market Incompleteness,"*Econometrica* 70(3), 907-928.

Kocherlakota, N. R. (1996): "Implications of Efficient Risk Sharing without Commitment," *Review of Economic Studies* 63 (4), 595-609.

Krueger, D. and F. Kubler (2008): "Markov Equilibria in Macroeconomics," in *The New Pelgrave Dictionary* ed. by S.N. Durlauf and L.E. Blume.

Kydland, F. E. and E. C. Prescott (1977): "Rules Rather than Discretion: The Inconsistency of Optimal Plans," *Journal of Political Economy* 85, 473–92.

Kydland, F. E. and E. C. Prescott (1980): "Dynamic Optimal Taxation, Rational Expectations and Optimal Control," *Journal of Dynamics and Contro*l 2, 79–91.

Ljungqvist L. and T. Sargent (2004): *Recursive Macroeconomic Theory.* The MIT Press.

Lucas, R. E. and N. Stokey (1983): "Optimal Fiscal and Monetary Policy," *Journal of Monetary Economics.*

Luenberger, D. G. (1969): *Optimization by Vector Space Methods.* New York: Wiley.

Mele A. (2010): "Repeated Moral Hazard and Recursive Lagrangeans," , Economics Department wp., Oxford University.

Marcet, A. (2008): "Recursive Contracts," in *The New Pelgrave Dictionary* ed. by S.N. Durlauf and L.E. Blume.

Marcet, A. and R. Marimon (1992): "Communication, Commitment and Growth," *Journal of Economic Theory* 58:2, 219–249.

Marcet, A. and R. Marimon (1998, 1999): "Recursive Contracts " *European University Institute, ECO 1998 #37 WP , Universitat Pompeu Fabra WP # 337.*

Marimon, R., Messner M. and N. Pavoni (2011): "Solving Recursive Contracts with Non-unique Solutions," *European University Institute.*

Marimon, R. and V. Quadrini (2011): "Competition, Human Capital and Income Inequality with Limited Commitment", *Journal of Economic Theory* (forthcoming)

Marimon, R. and J. Werner (2010): "On the Envelope Theorem without Differentiability," *European University Institute.*

Messner M. and N. Pavoni (2004): "On the Recursive Saddle Point Method," IGIER Working paper 255, Universita Bocconi.

Phelan, C. and R. M. Townsend (1992): "Computing the Optimal Insurance-Incentive Tradeoff for Multiperiod, Information Constrained Economies," Review of Economic Studies.

Phelan, C. and E. Stacchetti (2001):"Sequential Equilibria in a Ramsey Tax Model," *Econometrica* 69, 1491-518.

Rockafellar, R.T. (1970): *Convex Analysis.* Princeton, NJ: Princeton University Press..

Rockafellar, R.T. (1981): *The Theory of Subgradients and its Applications to Problems of Optimization. Convex and Nonconvex Functions.* Berlin: Heldermann Verlag,.

Rustichini, A. (1998a): "Dynamic Programing Solutions of Incentive Constrained Problems," *Journal of Economic Theory* 78(2), 329-354.

Rustichini, A. (1998b): "Lagrange Multipliers in Incentive-Constrained Problems," *Journal of Mathematical Economics,* 29, 365-380.

Stokey, N. L., R. E. Lucas and E. C. Prescott (1989): *Recursive Methods in Economic Dynamics.* Cambridge, Ma.: Harvard University Press.

Svensson, L. and N. Williams (2008): "Optimal Monetary Policy under Uncertainty: A Markov Jump-Linear-Quadratic-Approach," *Federal Reserve Bank of St. Louis Review.*

Thomas, J. and T. Worral (1990): "Income Fluctuations and Asymmetric Information: An Example of Repeated Principal-Agent Problem," *Journal of Economic Theory* 51: 367–390.

Solving recursive contracts with non-unique solutions

Ramon Marimon*
European University Institute and UPF - BarcelonaGSE

Matthias Messner
Bocconi University - IGIER

Nicola Pavoni
European University Institute and Bocconi University - IGIER

February 2011

Abstract

We extend Marcet and Marimon's recursive saddle-point method (1998, revised 2011) to encompass contracting problems with intertemporal incentive constraints with non-unique solutions. Our approach exploits the subdifferentiability properties of the corresponding value functions, which are convex and homogeneous in the co-state variable. We provide a unique selection procedure which guarantees that the co-state variable is a sufficient statistic for past Lagrange multipliers associated with the incentive constraints.

Keywords: Recursive saddle point, recursive contracts, dynamic programming, subgradients.
JEL code: C61, C63.

*I want to thank Jan Werner since our joint work has helped to shape the ideas of this paper.

1 Introduction

Marcet and Marimon (1998, revised 2011) (henceforth M&M) develop an elegant and flexible recursive saddle-point method, suitable for solving a wide class of dynamic optimization problems. Because of its tractability and computational advantages, throughout the last years many authors have started applying this approach. The list of papers which use it includes, among many others, Attanasio and Rios-Rull(2001), Ayagari, Marcet, Sargent and Seppälä (2002), Cooley, Marimon and Quadrini (2001), Friedman (1998), Kehoe and Perri (2002), Khan, King and Wolman (2000), Klein and Rios-Rull (2002), Marcet and Marimon (1992), Seppälä (2002), and Siu (2002). These applications include default models, optimal Ramsey taxation problems, and moral hazard models.

One of the key advantages of the M&M method with respect to dynamic programming is that the state space is not itself endogenous but it is predetermined, which is the usual sense of a recursive state. This can simplify the numerical analysis substantially with respect to the alternative approach –based on the work of Abreu, Pearce and Stacchetti (1990) – of using *promised utilities* as state variables, which often requires costly preliminary computations (see, for example, Abraham and Pavoni (2005) and Chang (1998)). The other key advantage of this method is that the state variables are simple objects. The dimensionality of the endogenous states only depends on the number of constraints. In contrast, in the usual techniques the dimensionality of endogenous states space depends on the number of shocks. In particular, when uncertainty is described by random variables with a continuous support one must maximize by choosing a function (an infinite dimensional object) even in the recursive formulation of the problem.

The 'recursive contract' approach of M&M builds on the duality between solutions to optimization problems ('Planner problems'; **PP**) and the saddle-points of the Langrangean associated with them (**SPP**). M&M essentially extend Bellman's principle of optimality to dynamic saddle-point problems, where the state space is enlarged with co-states acting as sufficient statistics of past Lagrange multipliers.

In duality theory, convexity and interiority assumptions guarantee that optimality problems have well defined Lagrangean saddle-points; the passage from **PP** to **SPP**. Furthermore, if Lagrangean multipliers exist, saddle-point solution controls are solutions to the original problem; that is, the sufficiency results –guaranteeing the passage from **SPP** to **PP**– are very general. In M&M (2011) **SPP** is a one-period saddle-point problem; that is, only the incentive constraints of the initial period are taken into account. M&M show how the standard duality equivalence (**PP** $\Leftrightarrow$ **SPP**) generalizes. However, the central element of their theory is to obtain a *saddle-point functional equation* (**SPFE**) generalizing Bellman's optimality principle to dynamic saddle-point problems and, ultimately, to determine the condition under which there is a duality equivalence: **PP** $\Leftrightarrow$ **SPFE**; which in practice requires a sequential formulation of **SPP** problems. They obtain such a sequential formulation by assuming *uniqueness* of solutions, or by making assumptions on primitives of the models that imply such uniqueness. While strictly concave return functions can guarantee the required *uniqueness*, for most problems it is a strong requirement (even if, once a solution has been found, may not be too difficult to verify whether is at least locally unique).

In fact, Messner and Pavoni (2004) provided an example where the lack of uniqueness resulted in dynamic saddle-point (**SPFE**) solutions which were not a solution to the original problem (**PP**) [1]. In this paper we revisit this example and we provide a natural extension of M&M (2011) as to encompass dynamic contracting problems without unique solutions. To this end, building on M&M (2001), we furhter exploit the subdifferentiability properties of the value functions, which are convex and homogeneous (of degree one) in the corresponding co-state variables. In particular, **SPFE** has non-unique solutions –which may not be solutions to the original problem **PP**- if, and only if, the value function is not differentiable in the co-state variable. Our third building block is the recent work on the *Envelope Theorem without differentiability* (Marimon and Werner (2010)), which allow us to make a proper selection in the case of multiple solutions. We first discuss the uniqueness problem.

2 The uniqueness problem

Before we revisit Messner and Pavoni (2004) example it is illustrative to see how 'the uniqueness problem' can already be seen in a simple static Pareto optimal problem.

Example 1. A simple Pareto problem

Consider the simple planner's problem of spliting an endowment among two agents, but one agent makes the offer subject to the participation constraints of both agents:

$$\begin{array}{rcl} \textbf{PP} & & \max_{a\in[0,y]} u_1(y-a) \\ \text{s.t.} \quad u_2(a) & \geq & \underline{u}_2 \end{array}$$

[1] Messner and Pavoni (2004) was in fact a 'counterexample' to M&M (1998), since in that version the *uniqueness* condition was not explicit.

Provided that there is an interior solution $\widetilde{a}$ and that $u_i' > 0$, $i = 1, 2$, this problem has a well defined saddle-point:

$$\min_{\gamma \geq 0} \max_{a \in [0,y]} \left\{ u_1(y-a) + \gamma \left(u_2(a) - \underline{u}_2 \right) \right\},$$

with a solution (a^*, γ^*), satisfying: $u_2(a^*) = \underline{u}_2$ and $\gamma^* = u_1'(y-a^*)/u_2'(a^*)$. As a result the 'value' of **PP** is given by

$$\begin{aligned} V(y) &= u_1(y-a^*) + \gamma^* \left(u_2(a^*) - \underline{u}_2 \right) \\ &= u_1(y-a^*) \end{aligned}$$

As standard maximization theory shows, when $u_i'' < 0$ $i = 1, 2$,we obtain the same unique solution a^* by solving the following Pareto problem:

$$\mathbf{PP}_\mu \quad \max_{a \in [0,y]} \left\{ u_1(y-a) + \mu u_2(a) \right\},$$

provided that $\mu = \gamma^*$; in other words, $\mu = \gamma^*$ is a sufficient statistic for **PP** when solving $\mathbf{PP}_\mu$.

However, when $u_i'' = 0$, $i = 1, 2$, $\gamma^* = 1$ and any $a^* \geq 0$ is a solution to $\mathbf{PP}_\mu$; therefore, with only one exception, the solutions to $\mathbf{PP}_\mu$ are not solutions to **PP.** This is the 'uniqueness problem': solving $\mathbf{PP}_\mu$ does not guarantee that we obtain a solution to the original problem **PP** when $\mathbf{PP}_\mu$ has not unique solutions .

To take a closer look at this problem, we fix y and consider the more general saddle-point problem $\mathbf{SPP}_\mu$:

$$W(\mu) = \min_{\gamma \geq 0} \max_{a \in [0,y]} \left\{ u_1(y-a) + \mu u_2(a) + \gamma \left(u_2(a) - \underline{u}_2 \right) \right\}. \tag{1}$$

Notice that, either in the non-linear case ($u_i'' < 0$ $i = 1, 2$) for any $\mu \geq 0$,or in the linear case ($u_i'' = 0$ $i = 1, 2$) for any $\mu \geq 0$, $\mu \neq 1$, W is differentiable, with

$$W'(\mu) = u_2(a^*(\mu))$$

where $(a^*(\mu), \gamma^*(\mu))$ is the unique solution to $\mathbf{SPP}_\mu$. In particular, in the linear case if $\mu < 1$, $\mu + \gamma^*(\mu) = 1$, $a^*(\mu) = u_2^{-1}(\underline{u}_2)$ and $W'(\mu) = \underline{u}_2$, while if $\mu > 1, \gamma^*(\mu) = 0$, $a^*(\mu) = y$ and $W'(\mu) = u_2(y)$ .However, in the linear case W is not differentiable at $\mu = 1$, and $\gamma^*(\mu) = 0$, while $a^*(\mu) \in [u_2^{-1}(\underline{u}_2), y]$. Nevertheless, since W is convex in μ, the right-hand and the left-hand derivatives of are well defined for any μ, and given by

$$\begin{aligned} W'^{+}(\mu) &= \lim_{t \searrow 0} \frac{W(\mu+t) - W(\mu)}{t} = \max_{a^* \in A^*(\mu)} \min_{\gamma^* \in \Lambda^*(\mu)} u_2(a^*(\mu)) \\ W'^{-}(\mu) &= \lim_{t \searrow 0} \frac{W(\mu-t) - W(\mu)}{t} = \min_{a^* \in A^*(\mu)} \max_{\gamma^* \in \Lambda^*(\mu)} u_2(a^*(\mu)), \end{aligned}$$

where $A^*(\mu) \times \Lambda^*(\mu)$ is the set of saddle-point solutions to (1) and the order of maximum and minimum does not matter[2]. In particular, in the linear case: $A^*(1) \times \Lambda^*(1) = [u_2^{-1}(\underline{u}_2), y] \times \{0\}$; that is, $W'^+(1) = u_2(y)$ and $W'^-(1) = \underline{u}_2$ are the extreme points of the subgradient set of W at $\mu = 1$, $\partial W(1)$. That is, a solution to $\mathbf{SPP}_\mu$ at $\mu = 1$ is a selection in $\partial W(1) = A^*(1) = [u_2^{-1}(\underline{u}_2), y]$ and, naturally, any selection results in the same value $W(1) = y$. However, if $\mathbf{PP}_\mu$ is the Pareto problem corresponding to the oiginal **PP,** only the selection corresponding to $W'^-(1)$ is the solution to the original **PP...** as if in order to make the appropriate selection one had to remember that in the original problem $\mu = 0$ and, therefore, 'the direction of arrival' to $\mu = 1$ was from the left!

While there is not much sense to talk about 'the direction of arrival' in a static problem this selection criterium becomes a natural one in a dynamic context; therefore, we turn to the dynamic example of Messner and Pavoni (2004)

Example 2. Messner and Pavoni (2004) .

We consider a dynamic genearlization of the previous example, concentrating the discussion in the linear case:

$$\begin{aligned} \mathbf{PP} \qquad & \sup_{\{a_t\}} \sum_{t=0}^{\infty} \beta^t (y - a_t) \\ \text{s.t. } a_t \in \ & [0, \overline{a}], \quad t \geq 0 \\ \sum_{n=0}^{\infty} \beta^n a_{t+n} \geq \ & b(1-\beta)^{-1}, t \geq 0 \end{aligned}$$

where $y > \overline{a} > b > 0, 0 < \beta < 1$ and $b < \beta\overline{a}$. Notice that, without loss of generality, we can add a participation constraint for the planner which will never be binding, such as

$$\sum_{n=0}^{\infty} \beta^n (y - a_t) \geq 0, t \geq 0.$$

In which case, we can rewrite the objective function as

$$\mathbf{PP}_\mu \qquad \sup_{\{a_t\}} \sum_{t=0}^{\infty} \beta^t \left[\mu^1 (y - a_t) + \mu^2 a_t\right],$$

where the original problem is given by $\mathbf{PP}_{\mu_0}$, with $\mu_0 = (1, 0)$**.**

The **SPFE** of this problems is given by (see Appendix for details):

$$\begin{aligned} & \mu^1(y - a^*) + \mu^2 a^* + \gamma^1 (y - a^*) + \gamma^2 (a^* - b(1-\beta)^{-1}) + \beta W(\mu + \gamma) \\ \geq \ & \mu^1(y - a^*) + \mu^2 a^* + \gamma^{1*} (y - a^*) + \gamma^{2*} (a^* - b(1-\beta)^{-1}) + \beta W(\mu + \gamma^*) \\ \geq \ & \mu^1(y - a) + \mu^2 a + \gamma^{1*} (y - a^*) + \gamma^{2*} (a - b(1-\beta)^{-1}) + \beta W(\mu + \gamma^*) \end{aligned}$$

[2] Here we follow Marimon and Werner (2010; Theorem 1), which generalizes Milgrom and Segal (2002; Theorem 5).

The value funtion of this problem satisfies

$$W(\mu) = \mu^1(y - a^*) + \gamma^{1*}(y - a^*) + \mu^2 a^* + \gamma^{2*}(a^* - b(1-\beta)^{-1}) + \beta W(\mu + \gamma^*) \quad (2)$$

where

$$\begin{aligned}
W(\mu) &= (\mu^1(y-b) + \mu^2 b)(1-\beta)^{-1}, \text{ if } \mu^2 \leq \mu^1 \quad (3)\\
W(\mu) &= y(1-\beta)^{-1}, \qquad \text{if } \mu^2 = \mu^1\\
W(\mu) &= (\mu^1(y-\overline{a}) + \mu^2\overline{a})(1-\beta)^{-1}, \text{ if } \mu^2 > \mu^1
\end{aligned}$$

and the *saddle-point policy correspondence* (*SP policy correspondence*) of **SPFE,** $(a^*, \gamma^*) \in \Psi(\mu)$ is

$$\begin{aligned}
\Psi(\mu) &= (b, (0, \mu^2 - \mu^1)), \text{ if } \mu^2 \leq \mu^1\\
\Psi(\mu) &= ([b, \overline{a}], (0,0)), \quad \text{if } \mu^2 = \mu^1\\
\Psi(\mu) &= (\overline{a}, (0,0)), \qquad \text{if } \mu^2 > \mu^1
\end{aligned}$$

Following M&M (2011) we can express (2) as

$$W(\mu) = \mu\omega^d(\mu) \equiv \sum_{j=1}^{2} \mu^j \omega_j^d(\mu)$$

where $\omega_j^d(\mu)$ is a partial directional derivative. Notice that $\omega^d(\mu) = \nabla W(\mu)$ whenever $\mu^2 \neq \mu^1$; i.e. $\omega^d(\mu)$ is uniquely defined, since W is differentiable in μ when $\mu^2 \neq \mu^1$. However, while $W(\mu)$ is uniquely defined when $\mu^2 = \mu^1$ its decomposition is not. As in Example 1, a selection from the *saddle-point policy correspondence* at $\mu^2 = \mu^1$, corresponds to a selection of the decomposition of $W(\mu)$ within the subgradient set:

$$\omega(\mu) \in \partial W(\mu) = [(y-b, b), (y-\overline{a}), \overline{a}]\,(1-\beta)^{-1}$$

In particular, when one starts with the orginal problem $\mathbf{PP}_{\mu_0}$, with $\mu_0 = (1,0)$, then $\gamma^*(\mu_0) = (0,1), a^*(\mu_0) = b$, and $\mu_1^2 = \mu_1^1$. However, if at $t = 1$ one chooses $a^*(\mu_1) \in (b, \overline{a}]$ and, consitently, maintains this choice for $t > 1$, then

$$(1-\beta)\omega^2(\mu_0) = b \neq a^* \in (1-\beta)\partial_2 W(\mu_1)$$

This example is a benchmark example in that W is differentiable in μ at $\mu_0 = (1,0)$. Furthermore, for any $\mu' = \mu + \gamma^*$ we know that $\gamma^* = \mu' - \mu$ is the 'direction of arrival' to μ'. We take this direction as the 'natural selection' criterium. The directional derivatives allows us to use a version of the Envelope Theorem (see Marimon and

Werner (2010) and Milgrom and Segal (2002)). Partial directional derivatives are well defined; for example, the second-partial left derivative is given by

$$\omega_2^-(\mu) \equiv \partial_2^- W(\mu) = \lim_{t \searrow 0} \frac{W(\mu - te^2) - W(\mu)}{t}, \tag{4}$$

where $e^2 = (0,1)$. In this example, at $\mu^2 < \mu^1$, $\omega_2^-(\mu) = \omega_2(\mu) = b(1-\beta)^{-1}$and at $\mu^2 = \mu^1$,

$$\omega_2^-(\mu) = \min\{z : z \in [b, \overline{a}]\}(1-\beta)^{-1} = b(1-\beta)^{-1}$$

Therefore, from $\mu_0 = (1,0)$ to $\mu_1 = \mu_0 + \gamma^*(\mu_0) = (1,1)$, the envelope in the direction $\gamma^*(\mu_0)$ results in the following individual recursive equation

$$\omega_2(\mu_0) = \min_{a^* \in A^*(\mu_0)} \max_{\gamma^* \in \Lambda^*(\mu_0)} \left\{a^* + \beta \partial_2^- W(\mu_0 + \gamma^*)\right\}, \tag{5}$$

where $A^*(\mu_0) \times \Lambda^*(\mu_0)$ is the set of saddle-point solutions to **SPFE** at μ_0; that is,

$$b(1-\beta)^{-1} = b + \beta b(1-\beta)^{-1},$$

As we see, in this example, with the 'direction of arrival' selection citerium (i.e. applying the Envelope Theorem in the direction $\gamma^* = \mu' - \mu$) solutions to **SPFE** at μ_0 are solutions to $\mathbf{PP}_{\mu_0}$ (see Appendix).

In the next section we provide the general formulation of our approach and show how the passage from **SPFE** to **PP** is guaranteed with this selection criterium. As our solution to this example suggests, this passage requires to keep 'the direction of arrival', γ^*, as part of the co-state variable.

3 The Dynamic Saddle-Point Problem

We follow M&M's (2011) notation and their characterization of the *Dynamic Saddle Point Problem*, associated with **SPFE.** In particular, they consider the following spaces of "value" functions

$$\begin{aligned}\mathcal{M}_b = \{&W : X \times \mathcal{R}_+^{l+1} \times S \to \mathcal{R} \\ &i) W(\cdot, \cdot, s) \text{ is continuous, and } W(\cdot, \mu, s) \text{ bounded, when } \|\mu\| \leq 1, \\ &ii) W(x, \cdot, s) \text{ is convex and homogeneous of degree one}\}\end{aligned}$$

and

$$\begin{aligned}\mathcal{M}_{bc} = \{&W \in \mathcal{M}_b \text{ and} \\ &iii) W(\cdot, \mu, s) \text{ is concave}\}.\end{aligned}$$

$\mathcal{M}_b$ is a space of continuous, bounded functions (in x), and convex and homogenous of degree one (in μ), while $\mathcal{M}_{bc}$ is the subspace of concave functions (in x). Both spaces are normed vector spaces with the norm

$$\|W\| = \sup\left\{|W(x,\mu,s)| : \|\mu\| \leq 1, x \in X, s \in S\right\}.$$

Whenever W satisfies (ii) it can be represented as $W(x,\,\mu,\,s) = \mu\omega(x,\,\cdot,\,s)$: therefore, it is convenient to define the corresponding spaces of functions:

$$\begin{aligned}
M_b = &\{\omega : X \times \mathcal{R}_+^{l+1} \times S \to \mathcal{R}^{l+1} \text{ s.t., for } j = 0, ..., l, \\
&i)\omega_j(\cdot,\,\cdot,\,s) \text{ is continuous, and } \omega_j(\cdot,\,\mu,\,s) \text{ bounded, when } \|\mu\| \leq 1 \\
&ii)\omega_j(x,\,\cdot,\,s) \text{ is convex and homogeneous of degree zero}\}
\end{aligned}$$

and

$$\begin{aligned}
M_{bc} = &\{\omega \in M_b \text{ s.t., for } j = 0, ..., l, \\
&iii)\omega_j(\cdot,\,\mu,\,s) \text{ is concave}\},
\end{aligned}$$

$\omega \in M$ uniquely defines a function $W \in \mathcal{M}$, given by $W \equiv \mu\omega$, but $W \in \mathcal{M}$ does not uniquely define a $\mathcal{R}^{l+1}$ valued function $\omega \in M$; it does when, in addition, W is differentiable in μ.

Given a function $\omega \in M$, and an initial condition (x,μ,s), the *Dynamic Saddle Point Problem* is given by

$$\begin{aligned}
&\textbf{DSPP} \qquad \inf_{\gamma\geq 0}\sup_{a}\left\{\mu h_0(x,\,a,\,s) + \gamma h_1(x,\,a,\,s) + \beta\ \mathrm{E}\left[\mu'\omega(x',\,\mu',\,s')|\,s\right]\right\} \\
&\text{s.t. } x' = \ell(x,\,a,\,s),\ p(x,a,\,s) \geq 0 \\
&\text{and } \mu' = \varphi(\mu,\,\gamma),
\end{aligned}$$

Here we assume that it has a solution[3] and therefore the following **SPFE** operator is well defined: $T^* : \mathcal{M} \longrightarrow \mathcal{M}$ given by

$$\begin{aligned}
&(T^*W)(x,\mu,s) = \min_{\gamma\geq 0}\max_{a}\left\{\mu h_0(x,\,a,\,s) + \gamma h_1(x,\,a,\,s) + \beta\ \mathrm{E}\left[W(x',\,\mu',\,s')|\,s\right]\right\} \\
&\text{s.t. } x' = \ell(x,\,a,\,s),\ p(x,a,\,s) \geq 0 \\
&\text{and } \mu' = \varphi(\mu,\,\gamma),
\end{aligned}$$

When $W \equiv \mu\omega$, with $\omega \in M$, and $\mathbf{DSPP}_{(x,\mu,s)}$ uniquely defines the values $h_0^j(x,\,a^*(x,\,\mu,\,s),\,s)$, $j = 0, ..., l$, then T^* defines a mapping, $T : M \longrightarrow M$, given by

[3]See M&M (2001 for more explicit assumptions guaranteeing the existence of a solution to **DSPP.**

$$(T\omega_j)(x,\, \mu,\, s) = h_0^j(x,\, a^*(x,\, \mu,\, s),\, s) + \beta\; \mathrm{E}\left[\omega_j(x^{*\prime}(x,\, \mu,\, s),\, \mu^{*\prime}(x,\, \mu,\, s),\, s')|\, s\right], \quad (6)$$

if $j = 0, ..., k$, and

$$(T\omega_j)(x,\, \mu,\, s) = h_0^j(x,\, a^*(x,\, \mu,\, s),\, s), \text{ if } j = k+1, ..., l.$$

M&M (2011) consider the case where the T map is well defined. We consider the more general case, where it may not be. To this end we extend the co-state with the 'direction of arrival', which acts as a recursive indicator of functions in M. More precisely, given $\gamma^* = \mu' - \mu$, let

$$d(\gamma^*) = sign\left(\frac{\mu'}{\|\mu'\|} - \frac{\mu}{\|\mu\|}\right),$$

where, given $z \in \mathcal{R}^{l+1}$, '$sign(z)$' is a vector of postive or negative ($+1$ or -1) indicators, which in our case determines the partial directional derivatives. More precisely, given $W \in \mathcal{M}$, and a directional vector d, we define (with a certain abuse of notation): $W^d(x, \mu, s) = W(x, \mu, s)$ but $W^d(x, \mu, s) \equiv \mu\omega^d(x, \mu, s) \equiv \mu\partial_d W(x, \mu, s)$; that is, the value is independent of the direction, d, but the decompostion is uniquely determined by the direction. We then consider the following operator:

$$(T^{'*}W)^d(x, \mu, s) = \min_{\gamma \geq 0} \max_a \left\{\mu h_0(x,\, a,\, s) + \gamma h_1(x,\, a,\, s) + \beta\; \mathrm{E}\left[W^{d'}(x',\, \mu',\, s')|\, s\right]\right\}$$

s.t. $x' = \ell(x,\, a,\, s),\ p(x, a,\, s) \geq 0$

and $\mu' = \varphi(\mu,\, \gamma),\ d' = d(\gamma^*)$ if $\gamma^* > 0$, $d' = d$ if $\gamma^* = 0$.

The choice of directional derivatives not only determines the *value decomposition representation* but also, consistent with it, the choice of maximizers; that is a^* is taken as the limiting value in the d' direction. With this selection the $T^{'*}$ operator uniquely defines a mapping, T', given by

$$(T'\omega_j^d)(x,\, \mu,\, s) = h_0^j(x,\, a^{d'*}(x,\, \mu,\, s),\, s) + \beta\; \mathrm{E}\left[\omega_j^{d'}(x^{d'*}\}(x,\, \mu,\, s),\, \mu^{*\prime}(x,\, \mu,\, s),\, s')|\, s\right],$$

if $j = 0, ..., k$, and

$$(T'\omega_j)(x,\, \mu,\, s) = h_0^j(x,\, a^{d'*}(x,\, \mu,\, s),\, s), \text{ if } j = k+1, ..., l.$$

The following lemma is the key step to our main result, which follows [4].

Lemma: If $T^* : \mathcal{M} \longrightarrow \mathcal{M}$ is a contraction mapping, with $W^* = T^*W^*$, then $W^{*\cdot} = T'^*W^{*\cdot}$ and $\omega^{*\cdot} = T'\omega^{*\cdot}$. Furthermore, for almost every μ, $\omega^{*\cdot} = \omega^*$; i.e. it is independent of d.

[4] We denote by $W^{*\cdot}$ the function W^{*d} for an arbitrary value d.

Proof: That $W^{*\cdot} = T'^{*}W^{*\cdot}$ follows from the fact that the $T^{'*}$ operator only adds to the T^{*} operator a selection criteria for the decompostion of W^{*}, which affects the choices, but not the contraction property of T'^{*}. Since this selection criterium uniquely determines the $T^{'*}$ map [[this will require a more detailed explanation] it follows that $\omega^{*\cdot} = T'\omega^{*\cdot}$. The last statement follows from the fact that W^{*} is convex in μ and finite, therefore non differentiable in at most countable many points (Rockafellar, 1970, Theorem 25.3).

Theorem: Assume the functions $h_0^j(x, \cdot, s)$, $j = 0, ..., l$, are invertible and that $T^{*} : \mathcal{M} \longrightarrow \mathcal{M}$ is a contraction mapping. If $(\mathbf{a}^{*}, \boldsymbol{\gamma}^{*})$ is generated by $W^{*\cdot}$, with $W^{*\cdot} = T'^{*}W^{*\cdot}$from (x, μ, s), then $(\mathbf{a}^{*}, \boldsymbol{\gamma}^{*})$ solves $\mathbf{SPP}_{\mu}$ at (x, s) and $\mathbf{a}^{*}$ is a solution to $\mathbf{PP}_{\mu}$ at (x, s).

Proof (sketch): The proof is basically the same than the one of Theorem 4 in M&M (2011), once it has been established that the $T^{'*}$ is uniquely determined by the selection criterium. As Theorem 5 in M&M (2011) stresses this was the missing piece to kill the 'counterexample'.

REFERENCES

ABRAHAM, A. AND N. PAVONI (2005): "Efficient Allocations with Moral Hazard and Hidden Borrowing and Lending"; *mimeo*, UCL and Duke University.

ABREU, D., AND E. STACCHETTI (1990): "Toward a Theory of Discounted Repeated Games with Imperfect Monitoring"; *Econometrica*, Vol. 58(5), 1041-1064.

ATTANASIO, O. AND J. V. RIOS-RULL (2000):"Consumption Smoothing in Island Economies: Can Public Insurance Reduce Welfare?" *European Economic Review*, Vol. 44, 1225-1258.

AYAGARI, R., A. MARCET, TH. J. SARGENT AND J. SEPPÄLÄ (2002): "Optimal Taxation without State-Contingent Debt"; *Journal of Political Economy*, Vol. 110(6), 1220-1254.

CHANG, R. (1998) "Credible Monetary Policy in an Infinite Horizon Model: Recursive Approaches," *Journal of Economic Theory,"* 81: 431-461.

COOLEY, T. F, R. MARIMON AND V. QUADRINI (2004): "Optimal Financial Contracts with Limited Enforceability and the Business Cycle," *Journal of Political Economy* 112, 817-47. FRIEDMAN, E. (1998): "Risk Sharing and the Dynamics of Inequality"; *mimeo.*

KEHOE, P. AND F. PERRI (2002): "International Business Cycles with Endogenous Incomplete Markets"; *Econometrica*, Vol. 70(3), 907-928.

KHAN, A., R. G. KING AND A. L. WOLMAN (2002): "Optimal Monetary Policy"; *Federal Reserve Bank of Philadelphia Working Paper* No. 01-5.

KLEIN, P. AND J. V. RIOS-RULL (2002): "Time-Consistent Optimal Fiscal Policy"; Forthcoming *International Economic Review.*

KOCHERLAKOTA, N. R. (1996): "Implications of Efficient Risk Sharing without Commitment," *Review of Economic Studies* 63 (4), 595-609. LIGON, E., J. P. THOMAS, AND T. WORALL(2000): "Mutual Insurance, Individual Savings and Limited Commitment" *Review of Economic Dynamics*, Vol. 3(2), 216-246.

LJUNGQVIST L. AND T. SARGENT (2004): *Recursive Macroeconomic Theory.* The MIT Press. LUENBERGER, D. G. (1969): "Optimization by Vector Space Methods"; New York: Wiley.

Fiscal and Monetary Policy in an Economy without Capital", *Journal of Monetary*

Economics. Vol. 12, 55-93.

Marcet, A. and R. Marimon (1992): “Communication, Commitment and Growth”; *Journal of Economic Theory*, Vol. 58, 853-881.

Marcet, A. and R. Marimon (1998, 1999, 2011): “Recursive Contracts ” *European University Institute, 2011 & ECO 1998 #37 WP , Universitat Pompeu Fabra WP # 337.* Messner, M. and N. Pavoni (2004): “On the Recursive Saddle Point Method”; *IGIER - Universita' Bocconi WP # 225.*

Phelan, Ch. and E. Stacchetti (2001): “Sequential Equilibria in a Ramsey Tax Model” *Econometrica* , Vol. 69(6), 1491-1518.

Phelan, Ch. and R. M. Townsend (1991): ”Computing Multi-Period, Information-Constrained Optima” *Review of Economic Studies* , Vol. 58, 853-881.

Rockafellar, R.T. (1981): *The Theory of Subgradients and its Applications to Problems of Optimization. Convex and Nonconvex Functions.* Berlin: Heldermann Verlag,. Seppälä, J. (1999): “Asset Prices and Business Cycles under Limited Commitment”; *mimeo.*

Siu, H. (2002): “Optimal fiscal and monetary policy with sticky prices”; *mimeo.*

Stockey, N. L., R. E. Lucas, with E. C. Prescott (1989): “Recursive Methods in Economics Dynamics”; Cambridge: Harvard University Press.