RECURSIVE CONTRACTS

Albert Marcet
Institut d'Anàlisi Econòmica CSIC, ICREA, UAB - Barcelona GSE, CEPR & MOVE

Ramon Marimon
European University Institute, UPF - Barcelona GSE, NBER and CEPR

This version: **March, 2017***

Abstract

We obtain a recursive formulation for a general class of optimization problems with *forward-looking* constraints which often arise in dynamic models, for example, in contracting problems with incentive constraints or in models of optimal policy. In this case, the solution does not satisfy the Bellman equation. Our approach consists of studying a recursive Lagrangian. Under standard general conditions there is a recursive *saddle-point* functional equation (analogous to a Bellman equation) that characterizes a recursive solution to the planner's problem. The recursive formulation is obtained after adding a co-state variable μ_t summarizing previous commitments reflected in past Lagrange multipliers. The time-consistent continuation solution is obtained by using the endogenous μ_t as the vector of weights in the objective function. Our approach is applicable to characterizing, and computing, solutions to a large class of dynamic contracting problems.

JEL classification: C61, C63, D58, E27

Key words: Recursive methods, dynamic optimization, Ramsey equilibrium, time inconsistency, limited commitment, limited enforcement, saddle-points, Lagrangian multipliers, Bellman equations.

*This is a substantially revised version of previously circulated papers with the same title (e.g. Marcet and Marimon 1998, 1999, 2011, 2016). We would like to thank Árpád Ábrahám, Fernando Alvarez, Truman Bewley, Charles Brendon, Edward Green, David Levine, Robert Lucas, Andreu Mas-Colell, Fabrizio Perri, Edward Prescott, Victor Ríos-Rull, Thomas Sargent, Robert Townsend and, especially, Fabian Schütze, Jan Werner and referees for comments over the years on this work, all the graduate students who have struggled through a theory in progress and, in particular, Matthias Mesner and Nicola Pavoni for pointing out a problem overlooked in the first versions. Support from MCyT-MEyC of Spain, CIRIT, Generalitat de Catalunya and the hospitality of the Federal Reserve Bank of Minneapolis are acknowledged. Marcet's research has been financed by the Axa Research Fund, European Research Council under the EU 7th Framework Programme (FP/2007-2013) Grant Agreement n. 324048 - APMPAL and the Programa de Excelencia del Banco de España.

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 (2012) describe a large number of applications to macroeconomic models. Under standard assumptions, the *Bellman equation* guarantees that the optimal solution has a recursive formulation. More precisely, it satisfies $a_t = \psi(x_t, s_t)$, where a_t denotes actions, s_t the exogenous shock to the economy and x_t is a small set of endogenous state variables. Importantly, ψ is a *time-invariant* policy function that solves the Bellman equation. We refer to this as the "standard dynamic programming" case. As is well known, in this case the solution is time consistent.

A key assumption needed to obtain the Bellman equation is that the feasible set for a_t is constrained only by (x_t, s_t). Unfortunately, many economic problems of interest include *forward-looking* constraints where future actions a_{t+j} also constrain the feasible set of a_t. This occurs, for example, in contracting problems where the principal chooses a contract subject to intertemporal participation constraints (see Example 1 below), and similarly in models of optimal policy design subject to intertemporal equilibrium constraints (see Example 2 below). Many dynamic games share the same feature.

In the presence of forward-looking constraints, optimal plans typically do not satisfy the Bellman equation and the solution does not have a standard recursive form. The reason is that contracting parties need to keep track of some additional variables summarizing commitments made in the past about today's choice. Otherwise, it leads to time-inconsistency, since agents will typically wish they could renege on their previous commitments. The absence of a standard recursive formulation greatly complicates the analysis and numerical solution of the model with commitment to promises.

In this paper, we provide an integrated approach for a recursive formulation of a large class of dynamic maximization problems with *forward-looking constraints*. We formulate a maximization problem $\mathbf{PP}_\mu$ where forward-looking constraints are embedded in the objective function. A contribution of the paper is to show that the optimal solution is obtained by solving at each point in time t a *continuation* planner problem $\mathbf{PP}_{\mu_t}$ (note that μ now has a subscript t) where the evolution of the weight μ_t is associated with the Lagrange multipliers of the forward-looking constraints.

We obtain a *saddle-point functional equation* (**SPFE**)**,** which is an analog of the *Bellman equation*, with the important difference that, while the *Bellman equation* solves a maximization problem, the *saddle-point functional equation* solves a saddle-point problem, as its name indicates. We then show *necessity*; that is, under standard general conditions, solutions to $\mathbf{PP}_\mu$ satisfy $a_t = \psi(x_t, \mu_t, s_t)$ for a *policy function* ψ, or a selection from a *policy correspondence* Ψ, which solves the **SPFE** with the weights μ following a pre-specified law of motion. We also provide a sufficiency condition guaranteeing that solutions to **SPFE** are solutions to $\mathbf{PP}_\mu$. This condition is satisfied if the value function W is differentiable in μ and, in particular, if the allocation solutions to the **SPFE** are locally unique (i.e. ψ is a *policy function*), as is the case in most economic applications.

The fact that our formulation is based on standard optimisation and dynamic programming tools facilitates the analysis and permits the application of a number of algorithms to obtain numerical solutions for dynamic stochastic models. For example, for a large class of models, accounting for *forward-looking constraints* translates into introducing time-varying Pareto weights into the objective function of $\mathbf{PP}_\mu$. More generally, the time-varying co-state μ_t enters as a *wedge* in the *stochastic discount factor* of $\mathbf{PP}_\mu$, showing the inter-temporal distortions due to the presence of *forward-looking constraints*.

We label $\mathbf{PP}_{\mu_t}$, given x_t as its initial condition, the *continuation problem*, because its solution coincides with the solution from period t onwards of the original problem $\mathbf{PP}_\mu$. Having this continuation problem at hand is at the core of the proof that the **SPFE** holds, and it facilitates the interpretation of time-inconsistent models. This continuation problem is also key to understanding some practical advantages of our approach. A commonly used tool for solving models with forward-looking constraints has been the promised-utility approach described in the pioneering works of Abreu, Pearce and Stacchetti (1990), Green (1987) and Thomas and Worrall (1988). In this approach, promised utilities need to be restricted so as to guarantee that the continuation problem is well defined. Computing the set of feasible utilities is often a major difficulty in this approach. One main advantage of our approach is that, under standard assumptions, the continuation problem $\mathbf{PP}_{\mu'}$ is guaranteed to have a solution for *any* $\mu' \geq 0$, sidestepping the computation of the set of feasible promised utilities. As we also discuss below, in many cases a recursive formulation in our approach is obtained with fewer decision variables and even fewer state variables than with promised utilities, allowing for a more efficient computation.

Our approach has already been used in many applications. A few examples are: growth and business cycles with possible default (Marcet and Marimon (1992), Kehoe and Perri (2002), Cooley, *et al.* (2004)); social insurance (Attanasio and Rios-Rull (2000)); optimal fiscal and monetary policy design with incomplete markets (Aiyagari, Marcet, Sargent and Seppälä (2002), Svensson and Williams (2008)); and political-economy models (Acemoglu, Golosov and Tsyvinskii (2011)). Furthermore, the introduction of the co-state variable μ_t to account for *forward-looking constraints* has proved to be a powerful instrument for analysing and comparing other economies with frictions (Chien, Cole and Lustig (2012)) and, in particular, in pricing contracts that endogenize *forward-looking constraints* or other frictions (Alvarez and Jermann (2000), Krueger, Perri and Lustig (2012)).

The main body of the theory is in Sections 4 and 5 of this paper (proofs are contained in the Appendix). Section 2 provides a basic introduction to our approach and Section 3 a couple of canonical examples. The relation to the promised utility approach is discussed in Section 3 and in more detail in Section 6. Section 7 concludes.

2 Formulating contracts as *recursive saddle-point problems*

In this section, we provide an outline of our approach to applying a recursive structure to a large class of models by extending dynamic programming methods to allow them to cover models with *forward-looking constraints*. We leave the formal results to sections 4 and 5. This section should be self-sufficient for a user of the method.

The class of models under study can be characterized as dynamic planning problems (**PP**) with a return function parameterized by a vector $\mu \in R^{l+1}$ as follows:

$$\mathbf{PP}_\mu : \quad V_\mu(x_0, s_0) = \sup_{\{a_t, x_t\}} \ \mathrm{E}_0 \sum_{j=0}^{l} \sum_{t=0}^{N_j} \beta^t \mu^j h_0^j(x_t,\, a_t,\, s_t) \tag{1}$$

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

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

given (x_0, s_0),

where ℓ, p, h_0, h_1 are known functions; β, x_0, s_0, μ are known constants or vectors, and $\{s_t\}_{t=0}^{\infty}$ an exogenous stochastic Markov process. Here, h_i^j is the j-th element of the function h_i for $i = 0, 1$, The solution is a *plan*[1] $\boldsymbol{a} \equiv \{a_t\}_{t=0}^{\infty}$, where $a_t(\ldots, s_{t-1}, s_t)$ is a state-contingent action.

The *forward-looking constraints* (3) are at the core of our analysis. We only consider $N_j = 0$ or ∞. Without loss of generality we assume $N_j = \infty$ for $j = 0, ..., k$, and $N_j = 0$ for $j = k+1, ..., l$ for a non-negative $k < l$. Note that this implies $N_0 = \infty$.

The case $N_j = \infty$ covers a large class of problems where discounted present values are part of the constraint, as in models with intertemporal participation constraints (see Example 1 in Section 3). Constraints with $N_j = 0$ cover cases where the intertemporal reactions of agents must be taken into account, as in dynamic Ramsey problems where agents' intertemporal Euler equations contain the actions of the government (see Example 2 in Section 3)[2].

Letting $\{a_t^*, x_t^*\}_{t=0}^{\infty}$ denote the solution of PP_μ, the value of the objective function at the optimum depends on (μ, x_0, s_0) and, therefore, we denote it by $V_\mu(x_0, s_0) \equiv E_0 \sum_{j=0}^{l} \sum_{t=0}^{N_j} \beta^t \mu^j h_0^j(x_t^*, a_t^*, s_t)$. Furthermore, note that h_0 appears both in the return function and the forward-looking part of the constraints (3). Models with an objective function of the form $E_0 \sum_{t=0}^{\infty} \beta^t r(x_t, a_t, s_t)$ for a time-invariant return function r are a special case of PP_μ if we take $h_0^0 = r$ and $\mu = (1, 0, ..., 0)$[3]. Considering a problem with a general μ as in (1) is useful since, as we show below, it delivers a continuation problem that characterizes a recursive solution.

Standard dynamic programming considers a special case of $\mathbf{PP}_\mu$ with two restrictions: *i)* constraints of the form (3) are absent or never binding, and *ii)* the objective function is a discounted infinite sum, i.e. $\mu_j = 0$ for $j > k$. As is well known[4], under fairly general assumptions, a necessary and sufficient condition for having a recursive formulation of $\mathbf{PP}_\mu$ is the existence of a value function

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

[2] Intermediate cases with finite $N_j > 0$ can be treated as a special case of $N_j = 0$. We discuss such a case at the end of Section 3.

[3] The return function r often does not appear in the forward-looking constraints. This is the case in Example 2 in Section 3. We can always make h_1^0 arbitrarily large so that (3) is not binding a.s. for $j = 0$ in order to accommodate this case under $\mathbf{PP}_\mu$.

[4] See, for example, Stokey, *et al.* (1989).

V_μ satisfying the following *Bellman functional equation*[5]:

$$V_\mu(x,s) \;=\; \sup_a \{\mu h_0(x,\,a,\,s) + \beta\,\mathrm{E}\,[V_\mu(x',s') \mid s]\} \qquad (4)$$
$$\text{s.t.}\quad p(x,a,\,s) \geq 0;\;\; x' = \ell(x,\,a,\,s').$$

Crucially, μ is the same on both sides of (4), while in the functional equation that we study the parameterization μ may change in the next period to μ' on the right-hand side; in fact, as we show, μ changes when the forward-looking constraints (3) are binding.

The power of dynamic programming is that, letting $\psi_\mu(x,s)$ be the arg max of the maximization problem in the 'Bellman equation', a sufficient condition for the optimal solution to $\mathbf{PP}_\mu$ is that it satisfies $a_t^* = \psi_\mu(x_t^*, s_t)$. Working with *time-independent* policy functions, as opposed to sequences, is of major help in characterizing and computing solutions to $\mathbf{PP}_\mu$. Furthermore, the solution is time consistent.

Unfortunately, as Kydland and Prescott (1977) pointed out, in the presence of forward-looking constraints (3) these dynamic programming results no longer hold, and the solution is often time *in*consistent.

An alternative functional equation

We show that a recursive formulation of the Lagrangian of this problem which provides the solution to $\mathbf{PP}_\mu$ can be achieved. Here we give an intuitive reasoning, while a formal proof is given in Sections 4 and 5.

Note that the Lagrangian of $\mathbf{PP}_\mu$ can be written as[6]

$$\mathcal{L}_\mu(\boldsymbol{a},\boldsymbol{\gamma}) = \mathrm{E}_0\Bigg[\sum_{j=0}^{l}\sum_{t=0}^{N_j}\beta^t\;\mu^j h_0^j(x_t,\,a_t,\,s_t) + \sum_{t=0}^{\infty}\sum_{j=0}^{l}\beta^t\gamma_t^j\;\mathrm{E}_t\sum_{n=1}^{N_j+1}\beta^t\Big(h_0^j(x_{t+n},\,a_{t+n},\,s_{t+n}) + h_1^j(x_t,\,a_t,\,s_t)\Big)\Bigg],$$

where γ_t is the Lagrange multiplier associated with (3). First, rearrange this Lagrangian using the law of iterated expectations to eliminate E_t in the *forward-looking* constraints; second, using Abel's summation to collect all the terms with arguments $(x_t,\,a_t,\,s_t)$, we find that for any $(\boldsymbol{a},\boldsymbol{\gamma})$ we can rewrite $\mathcal{L}_\mu$ as

$$\mathcal{L}_\mu(\boldsymbol{a},\boldsymbol{\gamma}) = \mathrm{E}_0\sum_{t=0}^{\infty}\beta^t\;\left[\mu_t\,h_0(x_t,\,a_t,\,s_t) + \gamma_t\,h_1(x_t,\,a_t,\,s_t)\right], \qquad (5)$$

where $\mu_{t+1} = \varphi(\mu_t,\,\gamma_t)$ for $\varphi : R_+^{l+1} \to R_+^{l+1}$ is given by

$$\begin{aligned}\varphi^j(\mu,\,\gamma) &\equiv \mu^j + \gamma^j \quad \text{for } j = 0,...,k\\ &\equiv \gamma^j \qquad\quad\;\; \text{for } j = k+1,...,l.\end{aligned} \qquad (6)$$

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

[6] Here we assume that the Lagrangian is well defined (e.g. Abel's summation can be applied) and ignore standard constraints (2) since they do not require any special treatment.

and with initial conditions $\mu_0 = \mu$.

Upon inspection of (5) and (6), it is 'intuitive' that $\mathcal{L}_\mu$ has a recursive structure similar to the programs amenable to dynamic programming; namely, the objective function (5) is a discounted sum with ime-invariant return functions (h_0, h_1), and past shocks enter into the transition function (6) and the return function at t only through the 'state variables' (x_t, μ_t). However, this interpretation relies on the fact that the Lagrangian takes $(\boldsymbol{a}, \boldsymbol{\gamma})$ as sequences of decision variables *and* on the introduction of $\boldsymbol{\mu} \equiv \{\mu_t\}_{t=0}^{\infty}$ through (6) as a co-state variable. More precisely, $\mathcal{L}_\mu$ has a recursive structure when μ is treated as a state variable, with the transition $\mathcal{L}_\mu \longrightarrow \mathcal{L}_{\mu'}$ given by (6).

The proof that this recursive formulation provides the optimal solution to PP_μ is not a straightforward application of the Bellman equation to (5)-(6). Whilst the Bellman equation solves a functional equation through its use of the *maximum* operator, the Lagrangian approach, given by (5)-(6), defines a *saddle-point problem* The goal of this paper is to show that a functional equation analogous to Bellman's can be derived for saddle-point problems, and that this provides a convenient recursive representation of the solution to $\mathbf{PP}_\mu$.

We now introduce **notation for saddle point problems** and define a functional operator analogous to Bellman's. Given a function $\mathcal{F} : Y \times Z \longrightarrow R$, using the standard definition of a *saddle-point* of $\mathcal{F}$ as a point $(y^*, z^*) \subset Y \times Z$ such that

$$\mathcal{F}(y^*, z) \geq \mathcal{F}(y^*, z^*) \geq \mathcal{F}(y, z^*), \text{ for any } z \in Z \text{ and } y \in Y, \tag{7}$$

we call the problem of finding such an (y^*, z^*) a saddle-point problem, which we denote as

$$\underset{\inf_z, \sup_y}{\mathrm{SP}} \mathcal{F}(y, z),$$

and the *saddle-point* solving this problem we denoted as

$$(y^*, z^*) \equiv \arg \underset{\inf_z, \sup_y}{\mathrm{SP}} \mathcal{F}(y, z)$$

Note that in this definition the subindeces inf and sup in SP only denote which variables are minimised or maximised, i.e. which variables are on the right or the left side in the string of inequalities (7). Therefore there is no ordering or sequentiality of the inf and sup operators: a saddle point satisfies both inequalities in (7) simultaneously[7].

We generalize Bellman's dynamic programming maximization to this class of saddle-point problems. In particular, under standard convexity (and interiority) assumptions, if $\boldsymbol{a}^*$ is a solution to $\mathbf{PP}_\mu$

[7] In previous versions of this paper, even though we were explicitly considering only saddle-points, we have denoted the saddle-point problem as $\inf_{z \in Z} \sup_{y \in Y} F(y, z)$. Changing this to $\mathrm{SP}_{\inf_z \sup_y} \mathcal{F}(y, z)$ is a purely notational issue and, obviously, it does not affect the results. Apparently, the previous terminology was confusing as it lead some readers to believe that we were considering sequential problems of the form $\inf_{z \in Z} \left[\sup_{y \in Y} \mathcal{F}(y, z)\right]$, where first the sup over y, given z, is considered and then the inf is found over z, given the solution of the sup problem (Stackelberg games have this form). As is well known the inf [sup] and sup[inf] sequential problems can have different solutions and, in that case, the saddle-point may not exist (see, for example, Bertsekas (2009)). The current notation should avoid any confusion, since it highlights that we do not consider sequential problems or the *duality* between the inf [sup] and sup[inf] problems (more on this in Footnote 26).

at (x, s), then there are Lagrange multipliers $\boldsymbol{\gamma}^*$ such that $(\boldsymbol{a}^*, \boldsymbol{\gamma}^*)$ is a *saddle-point*[8]. Furthermore, if we let $W(x, \mu, s) = \mathrm{SP}_{\inf_{\boldsymbol{\gamma}}, \sup_{\boldsymbol{a}}} \mathcal{L}_\mu(\boldsymbol{a}, \boldsymbol{\gamma})$, then W satisfies the following *saddle-point functional equation:*

$$\textbf{SPFE} \quad W(x, \mu, s) = \underset{\inf_{\gamma \geq 0}, \sup_a}{\mathrm{SP}} \{\mu h_0(x, a, s) + \gamma h_1(x, a, s) + \beta \, \mathrm{E}\left[W(x', \mu', s') | \, s\right]\} \tag{8}$$
$$\text{s.t. } x' = \ell(x, a, s'), \; p(x, a, s) \geq 0$$
$$\text{and } \mu' = \varphi(\mu, \gamma).$$

While (8) formally generalizes the Bellman equation, there are three differences between (8) and the standard Bellman equation (4): *i)* it is a *saddle-point problem* rather than a maximization problem; *ii)* μ is an argument of the value function W, and *iii)* the law of motion for μ is added as a constraint.

We show the necessity of **SPFE**. This implies that the solution to $\mathbf{PP}_\mu$ indeed satisfies $(a_t^*, \gamma_t^*) = \psi(x_t^*, \mu_t^*, s_t)$ for a time-invariant function $\psi(x, \mu, s)$ which is the *saddle-point* that solves the **SPFE** and γ_t^* is the Lagrange multiplier vector of constraints (3). Furthermore, the value function of $\mathbf{PP}_\mu$ solves this functional equation, namely $W(x, \mu, s) = V_\mu(x, s)$.

We also show the sufficiency of **SPFE** under fairly general conditions. If $(\boldsymbol{a}^*, \boldsymbol{\gamma}^*) \equiv \{a_t^*, x_t^*\}_0^\infty$ is obtained from a selection of a *policy correspondence* of a value function satisfying (8) for all (x, μ, s), and a consistency condition is satisfied, then $(\boldsymbol{a}^*, \boldsymbol{\gamma}^*) \equiv \{a_t^*, x_t^*\}_0^\infty$ is the solution of $\mathbf{PP}_\mu$ at (x, s). The proof of this result (Theorem 4) is a little more involved, since it must be shown that $(\boldsymbol{a}^*, \boldsymbol{\gamma}^*)$ properly accounts for the original forward-looking constraints (3), which are not present in (8).

In sum, from the user's perspective, the main thing to retain is that a recursive solution is obtained by adding a co-state variable μ that is a function of the Lagrange multiplier of the forward-looking constraints in previous periods. As seen from (6), this state variable is the sum of past multipliers $\mu_{t+1}^{j,*} = \sum_{k=0}^{t} \gamma_k^{j,*} + \mu^j$, for $j \leq k$, namely when $N_j = \infty$ (i.e. constraints involving discounted sums), and it is the past multiplier $\mu_{t+1}^{j,*} = \gamma_t^{j,*}$ for $j > k$, namely when $N_j = 0$ (i.e. constraints involving one future period).

The time-inconsistency problem and the advantages of our approach

Assuming that the exogenous stochastic process is Markovian, the standard Bellman equation for maximization problems defines a policy function, for example, ψ_μ in (4) defines $(x_0, s_0) \to a_0^*$. In the standard dynamic programming case this policy function is time-consistent: reoptimization at the new state is also a continuation solution from the original state, formally, $\psi_\mu(\ell(x_0, a_0^*, \cdot), \cdot) = a_1^*(\cdot)$, where a_1^* is the optimal contingent solution given x_0. However, as is well known, in the presence of *binding* forward-looking constraints (3) this does not hold: reoptimization of $\mathbf{PP}_\mu$ at (x_1^*, s_1) would not yield a_1^*.

As mentioned above, with our approach the key is that if one optimizes $\mathbf{PP}_{\mu_1^*}$ (note the subscript is now μ_1^*) with initial conditions (x_1^*, s_1) the solution coincides with the continuation of the original solution $\{a_t^*, x_t^*\}_{t=1}^\infty$. This means that $\mathbf{PP}_{\mu_1^*}$ is a continuation problem in our approach. In cases when $\mu_1^* \neq \mu$ (i.e. $\gamma_0^* \neq 0$), solving $\mathbf{PP}_\mu$ at (x_1^*, s_1) would be time-inconsistent.

[8] We abstract here from the fact that these assumptions – in particular, interiority – require specific topologies if one considers the *infinite-dimensional saddle-point* (see Section 4 and Appendix). In the rest of the paper we say "$\mathbf{PP}_\mu$ at (x, s)" referring to the problem $\mathbf{PP}_\mu$ for given constants μ and initial conditions $(x_0, s_0) = (x, s)$.

The transition $\mathbf{PP}_{\mu_t^*} \to \mathbf{PP}_{\mu_{t+1}^*}$ captures several advantages of our approach. First, we use it as a step in proving that the necessity of **SPFE** holds. Second, it clarifies why we have one key advantage over the promised utility approach: under mild standard assumptions this continuation problem has a solution for *all* μ. Therefore, the only constraint on this co-state variable is that $\mu_t \in R_+^{l+1}$, while, as we will see, co-state variables in the promised utility approach need to be constrained appropriately and this adds many complications to the implementation of that approach. Third, as already mentioned, $\mathbf{PP}_{\mu_t^*}$ provides a natural way to check for time consistency: the solution to $\mathbf{PP}_{\mu}$ is time-consistent when its objective function coincides with (or is proportional to) the objective function of $\mathbf{PP}_{\mu_1^*}$. Fourth, our approach provides a useful economic intuition about how to design optimal contracts (institutions or mechanisms) subject to intertemporal incentive constraints and on how to 'price' the costs of these constraints, in order to decentralise these contracts.

3 Two Examples

In this Section we illustrate our approach with two examples. In the first, there are only *intertemporal participation constraints*, so it is a case where $N_j = \infty$ for all j (i.e. $k = l$); in the second, there is only one *intertemporal one-period (Euler) constraint* and hence it is a case with $l = 1$, $N_1 = 0$ (i.e. $k = 0$). The first is similar to the model studied in Marcet and Marimon (1992), Kocherlakota (1996), and 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 et al. (2002) and it is a canonical model with Euler constraints, as in Ramsey equilibria in optimal fiscal and monetary policy.

3.1 Example 1: Intertemporal participation constraints.

We consider a model of a partnership, where several agents can share their individual risks and jointly invest in a project which cannot be undertaken by single (or subgroups of) agents. There is a single consumption good and l infinitely-lived consumers. The preferences of agent j are represented by $E_0 \sum_{t=0}^{\infty} \beta^t \ u(c_t^j)$; u is assumed to be bounded, strictly concave and monotone; c represents individual consumption. Agent j receives an endowment of consumption good y_t^j at time t and $y_t = (y_t^1, \ldots, y_t^l)$. Agent j has an outside option that delivers total utility $v_j^a(y_t)$ if he leaves the contract in period t, where v_j^a is some known function. It is often assumed that the outside option is the autarkic solution where agent j consumes only his endowment from t onwards. In that case $v_j^a(y_t) = \mathrm{E}\left[\sum_{n=0}^{\infty} \beta^n u(y_{t+n}^j) \mid y_t\right]$. This implicitly assumes that if agent j defaults in period t he is permanently excluded from the partnership so he has no further claims on its production or capital in, and after, period t.

Total production is given by $F(k, \theta)$, where k is capital and θ a productivity shock. Production can be split into consumption c and investment i; capital depreciates at the rate δ. The process $\{\theta_t, y_t\}_{t=0}^{\infty}$ is assumed to be jointly Markovian and the initial conditions (k_0, θ_0, y_0) are given. The planner looks for Pareto optimal allocations that ensure that no agent ever leaves the contract, so it has to be ensured that the future discounted utility of all agents at any period and state is at least as high as their outside option $v_j^a(y_t)$. The planner's problem takes the form:

$$\max_{\{c_t, i_t\}} \quad \mathrm{E}_0 \sum_{t=0}^{\infty} \beta^t \sum_{j=1}^{l} \alpha^j \; u(c_t^j) \tag{9}$$
$$\text{s.t. } k_{t+1} = (1-\delta)k_t + i_t,$$
$$F(k_t, \theta_t) + \sum_{j=1}^{l} y_t^j \geq \sum_{j=1}^{l} c_t^j + i_t, \text{ and}$$
$$\mathrm{E}_t \sum_{n=0}^{\infty} \beta^n \; u(c_{t+n}^j) \geq v_j^a(y_t) \qquad \text{for all } j = 1, ..., l \text{ and } t \geq 0.$$

It is easy to map this planner's problem into our $\mathbf{PP}_\mu$ formulation if we take $\mu = \alpha$, $s \equiv (\theta, y)$; $x \equiv k$; $a \equiv (i, c)$; $\ell(x, a, s) \equiv (1-\delta)k+i$; $p(x, a, s) \equiv F(k,\theta)+\sum_{j=1}^{l} y^j - \left(\sum_{j=1}^{l} c^j + i\right)$; $h_0^j(x, a, s) \equiv u(c^j)$; $h_1^j(x, a, s) \equiv u(c^j) - v_j^a(y_t)$, $j = 1, ..., l$[9].

The Lagrangean $\mathcal{L}_\mu$ can be found to be

$$\mathcal{L}_\mu = \mathrm{E}_0 \sum_{t=0}^{\infty} \beta^t \sum_{j=1}^{l} \left[\mu_{t+1}^j \; u(c_t^j) - \gamma_t^j \, v_j^a(y_t)\right],$$

for $\mu_{t+1} = \mu_t + \gamma_t$ with initial conditions $\mu_0^j = \alpha^j$ $j = 1, ..., l$, and for feasible consumption allocations.

The **SPFE** takes the form

$$W(k,\, \mu, y, \theta) \; = \underset{\inf_{\gamma \geq 0}, \sup_{c,i}}{\mathrm{SP}} \left\{ \sum_{j=1}^{l} \left[\left(\mu^j + \gamma^j\right) u(c^j) - \gamma^j v_j^a(y)\right] \right.$$
$$\left. + \beta \; \mathrm{E}\left[W(k',\, \mu', y', \theta') \,|y, \theta\,\right] \right\}$$
$$\text{s.t. } k' = (1-\delta)k + i, \quad F(k,\theta) + \sum_{j=1}^{l} y^j \geq \sum_{j=1}^{l} c^j + i$$
$$\text{and } \mu' = \mu + \gamma.$$

Our results below guarantee that $W(k,\, \mu, y, \theta) = V_\mu(k, y, \theta)$ solves this functional equation and that letting ψ be the 'policy function' delivering the *saddle-points* of the right-hand side of the **SPFE**, the solution to the problem of interest (9) satisfies

$$\begin{aligned} (\gamma_t^*, c_t^*, i_t^*) &= \psi(k_t^*,\, \mu_t^*,\, \theta_t,\, y_t) \text{ and} \\ \mu_{t+1}^* &= \mu_t^* + \gamma_t^*, \end{aligned} \tag{10}$$

with initial conditions $(k_0, \mu_0,\, \theta_0,\, y_0)$, where $\mu_0 = (0, \alpha)$.

The fact that $\mathrm{PP}_{\mu_1^*}$ is the continuation problem tells us that the solution after period $t = 1$ coincides with the solution to (9) given initial conditions $(k_1^*, \theta_1,\, y_1)$ provided that the weights α of the agents in the objective function of (9) are replaced by $\mu_1^* = \alpha + \gamma_0^*$. Therefore, the co-state variables μ_t^* are the weights that the planner "should" assign to each agent, instead of the initial weights α, if

[9]Note that we simply eliminate $j = 0$.

the planner were to recover the initial solution by solving the model given state variables at $t = 1$. The variable μ_1^* is all that needs to be remembered from the past in state (k_1^*, θ_1, y_1).

Since, with standard assumptions, a solution to the continuation problem $\text{PP}_{\mu_1^*}$ exists for any $\mu_1 \in R_+^{l+1}$, in our approach we completely sidestep the complication of having to compute the set of feasible continuation promised utilities as would happen with the promised-utility approach.

This recursive formulation helps to characterize the solution: the weights μ_t^* evolve according to whether or not agents' participation constraints are binding. Every time that the participation constraint for an agent is binding, his weight is permanently increased by the amount of the corresponding Lagrange multiplier, although his relative weight evolves according to: $\mu_t^j / \sum_{i=1}^l \mu_t^i$. An agent is induced not to default by increasing his consumption not only in the period where he is tempted to default (i.e. when his participation constraint is binding) but, if possible, permanently and therefore smoothly over time. If only one agent is ever tempted to default (one-sided limited commitment as in Marcet and Marimon (1992)) his share of total consumption will permanently increase when his outside option is binding. With multi-sided limited commitment, his share will decrease when other agents need to be prevented from defaulting.

More precisely, due to these changing weights, relative marginal utilities across agents are not constant when participation constraints are binding, since the first-order-conditions imply

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

It follows that individual paths of consumption depend on individual histories, in particular, on past 'temptations to default' γ_{t-j}, and not just on the initial wealth distribution and the aggregate consumption path as in Arrow-Debreu competitive allocations. That is, individual consumption does not *co-move* perfectly with current aggregate consumption. It potentially depends on all past shocks $\{y_n, \theta_n\}_{n=0}^{t-1}$, but this dependence on the past is completely summarized by μ_t. If enforcement constraints are never binding (e.g. punishments are severe enough) then $\mu_t = \alpha$ and we recover the "constancy of the marginal utility of expenditure", and in that case individual consumptions are a fixed proportion of current aggregate consumption. In other words, the evolution of the co-state variables can also be interpreted as the evolution of the distribution of wealth. If intertemporal participation constraints are binding infinitely often, then there is a non-degenerate distribution of consumption in the long-run, in contrast to an economy where intertemporal participation constraints cease to be binding, as in an economy with full enforcement[10]. The evolution of the weights μ also helps in analyzing the price decentralization of contracts and characterizing the capital accumulation process[11].

The intertemporal Euler equation of $\mathbf{PP}_\mu$ at t, is given by:

$$\mu_{t+1}^j u'(c_t^j) = \beta \; \text{E}_t \left[\mu_{t+2}^j u'(c_{t+1}^j)(F_{k_{t+1}} + 1 - \delta) \right]. \tag{11}$$

[10]See, for example, Broer (2013) for a characterization of the non-degenerate stationary distribution of consumption, in a similar model with a finite number of types and a continuum of agents of each type.

[11]For a more detailed analysis of price decentralization in economies with limited enforcement, see Alvarez and Jermann (2000), Kehoe and Perri (2002) and Krueger, Lustig and Perri (2008).

In the first best allocation this equation holds for constant $\mu^j = \alpha^j$, for all j and t. The presence of time-varying μ in this equation shows how *limited enforcement constraints* introduce a *wedge* in agents' *stochastic discount factors*: $\beta \frac{\mu^j_{t+2} u'(c^j_{t+1})}{\mu^j_{t+1} u'(c^j_t)}$ – that is, it shows how these constraints distort consumption allocations and, therefore, prices, when the planner's problem is decentralised.

In order to find numerical solutions, one has to find a policy function ψ of the form (10) such that (11) and the participation and the feasibility constraints hold for all periods. Indeed, for parameterizations where the participation constraints are binding infinitely often we could have $\mu_t \to \infty$. Therefore, it is important to renormalize the vector μ_t – for example, with $\widehat{\mu}^j_t = \mu^j_t / \sum_{j=1}^J \mu^j_t$.

It should also be noted that the value function of **SPFE** takes the form $W(k,\, \alpha, y, \theta) = \sum_{j=1}^l \alpha^j \omega_j(k,\, \alpha, y, \theta) = \sum_{j=1}^l \alpha^j \mathrm{E}_0 \sum_{t=0}^\infty \beta^t \; u(c^{*j}_t)$ and, along the solution path, "individual values" satisfy the following recursive equations:

$$\omega_j(k^*_t,\, \mu^*_t, y_t, \theta_t) = u(c^{*j}_t) + \beta\; \mathrm{E}\left[\omega_j(k^*_{t+1},\, \overset{*}{\mu}_{t+1}, y_{t+1}, \theta_{t+1})\, | y_t, \theta_t\right]. \tag{12}$$

In contrast with the promised-utility approach, equations (12) are part of the solution to the **SPFE**. Therefore these equations are not 'promise-keeping' constraints added to the $\mathbf{PP}_\mu$ (see Section 4); they just define the individual discounted utilities.

3.2 Example 2: Intertemporal one-period constraints: a Ramsey problem

We present an abridged version of the optimal taxation problem under incomplete markets studied by Aiyagari at al. (2002). A representative consumer solves

$$\max_{\{c_t, e_t, b_t\}} E_0 \sum_{t=0}^\infty \beta^t \; [u(c_t) + v(e_t)]$$

$$\text{s.t.} \qquad c_t + b_{t+1} p^b_t = e_t(1 - \tau_t) + b_t,$$

for a given b_0, where c is consumption and e is effort (e.g. hours worked). The government must finance exogenous random expenditures g by setting tax rates τ, issuing real riskless bonds b and fully committing to future tax rates. The process $\{g_t\}_{t=0}^\infty$ is Markovian. Feasible allocations satisfy $c_t + g_t = e_t$. The bond and labor markets are competitive and $(g_0, ..., g_t)$ is public information at t. The government's budget mirrors that of the representative agent [12].

In a Ramsey equilibrium, the government chooses sequences of taxes and debt that maximize the utility of the consumer subject to the allocations being a competitive equilibrium. Substituting the equilibrium equations into the budget constraint of the consumer, the Ramsey equilibrium can be found by solving[13].

[12] As usual, Ponzi games need to be ruled out. This can be done with a natural debt limit that is not binding in equilibrium.

[13] As explained in Aiyagari et al., (2002) the inequality in the budget constraint is justified by the assumption that, if b_t is large and negative, the government distributes excess returns from savings via a lump sum transfer.

$$\max_{\{c_t,b_t\}} E_0 \sum_{t=0}^{\infty} \beta^t \; [u(c_t) + v(e_t)]$$
$$\text{s.t.} \qquad E_t \left[\beta b_{t+1} u'(c_{t+1})\right] \geq u'(c_t)(b_t - c_t) - e_t v'(e_t), \tag{13}$$

for a given b_0.

This problem is a special case of $\mathbf{PP}_\mu$ when we take $\mu = (1,0)$, $s \equiv g$; $x \equiv b$, $a \equiv (c, b')$, $\ell(x, a, s') \equiv b'$, $h_0^0(x, a, s') \equiv u(c) + v(e)$, $h_0^1(x, a, s') \equiv bu'(c)$, $h_1^1(x, a, s') \equiv u'(c)(c - b) + ev'(e)$, $N_1 = 0$ and a very large h_1^0 ensuring that $\gamma_t^0 = 0$ a.s.

Using $\mu_t^0 = 1$ for all t and for $\mu_0^1 = 0$, the objective function of the Lagrangian (5) becomes

$$\mathcal{L}_\mu = \mathrm{E}_0 \sum_{t=0}^{\infty} \beta^t \; \left[\mu_t^0 \left(u(c_t) + v(e_t)\right) + \mu_t^1 b_t u'(c_t) + \gamma_t^1 \left[u'(c_t)(c_t - b_t) + e_t v'(e_t)\right]\right]. \tag{14}$$

The **SPFE** takes the form

$$\begin{aligned} W(b,\, \mu, g) = & \underset{\inf_{\gamma^1 \geq 0}, \sup_{c,b'}}{\mathrm{SP}} \{\mu^0 \left[u(c) + v(e)\right] + \mu^1 b u'(c) \\ & + \gamma^1 \left[u'(c)(c - b) + ev'(e)\right] \;\; + \beta \; \mathrm{E}\left[W(b',\, \mu',\; g')\, |g\,\right]\} \\ \text{s.t.} \quad & \mu^{0\prime} = \mu^0, \;\, \mu^{1\prime} = \gamma^1. \end{aligned}$$

Letting ψ be the policy function defined by the *arginfsup* of the above saddle-point problem, efficient allocations satisfy

$$(c_t^*, b_{t+1}^*, \gamma_t^{1*}) \; = \psi(b_t^*,\, \mu_t^*,\, g_t), \tag{15}$$

for $\mu_{t+1}^* = (1, \gamma_t^{1*})$ with initial conditions $(b_0, \mu_0,\, g_0)$, where $\mu_0 = (1, 0)$.

The objective function of the continuation problem at t $\mathbf{PP}_{\mu_t^*}$ is given by

$$E_t \sum_{j=0}^{\infty} \beta^j \; [u(c_{t+j}) + v(e_{t+j})] + \mu_t^{1*} b_t^* u'(c_t). \tag{16}$$

The term $\mu_t^{1*} b_t^* u'(c_t) = \gamma_{t-1}^{1*} b_t \; u'(c_t)$ in (16) captures the commitment to choose taxes at t that guarantee the Euler equation (13) at $t - 1$. These terms are added to obtain an optimal manipulation of the interest rate so as to minimize the cost of debt in the presence of unexpected shocks[14]. The term $\gamma_t^1 \left(u'(c_t)c_t + e_t v'(e_t)\right)$ in (14) is also present in a complete markets version of the model but with a constant γ^1; the term $\left(\mu_t^1 - \gamma_t^1\right) b_t u'(c_t) = \left(\gamma_{t-1}^1 - \gamma_t^1\right) b_t u'(c_t)$ is the additional term due to incomplete markets.

It is clear that the solution of the continuation problem $\mathrm{PP}_{\mu_1^*}$ exists for any γ_0^{1*}, since the above objective function is continuous and bounded above under standard conditions. As in Example 1, we completely sidestep the complication of having to compute the set of feasible continuation values, as any $\gamma_0^1 \in R_+$ is feasible.

[14]See Faraglia, Marcet and Scott (2014) for a detailed description of interest rate manipulation through tax policy under full commitment.

The first-order conditions of the Ramsey problem imply that solutions satisfy

$$\mathrm{E}_t\left[(\gamma_t^1 - \gamma_{t+1}^1)u'(c_{t+1})\right] = 0. \tag{17}$$

Aiyagari et al. (2002) show that this implies that optimal fiscal policy under incomplete markets consists of modifying the deadweight loss of taxation γ_t^{1*} in each period and ensuring that it follows a risk-adjusted martingale satisfying (17).

A numerical solution involves finding a policy function (15) such that the FOC (17), the budget constraint (13) and feasibility hold (approximately). As we noted, in principle, γ_t^{*1} is unbounded, which can create problems in computing solutions, but for $\gamma_t^{*1} > 0$ one can renormalize $\overset{*}{\mu}_{t+1} = (1, \gamma_t^{*1})$ to $\widehat{\overset{*}{\mu}}_{t+1} = (1/\gamma_t^{*1}, 1)$. Therefore, for very large values of $\overset{*}{\mu}_t = \gamma_{t-1}^{1*}$, $\mathrm{PP}_{\mu_t^*}$ can be approximated by PP_μ with $\mu = (0, 1)$ – that is, with an objective function proportional to $u'(c_t)$ (since b_t^* is predetermined). With such an approximation, one only needs to approximate ψ when γ_t^{*1} belongs to a compact set $[0, Q]$ and an accurate non-linear approximation can be achieved.

3.3 Recursive Lagrangian vs. Promised-utility

The promised-utility and our approach provide recursive characterizations of the solution to $\mathbf{PP}_\mu$. In our approach the co-state variable is a vector μ satisfying a simple exogenous constraint: $\mu \in R_+^{l+1}$, while in the promised-utility approach, it is a vector – say, ω – which must satisfy an endogenous 'promise-keeping' constraint. Obviously, both approaches provide the same solutions $\{a_t^*, x_t^*\}$, but they are conceptually and practically quite different.

A practical difference arises in characterizing solutions. In particular, our co-state variable $\overset{*}{\mu}_t$ explicitly shows the 'wedge' between a constrained and an unconstrained solution, which can be used to derive prices to decentralize the constrained-efficient solution (as in Example 1). $\overset{*}{\mu}_t$ also explicitly shows the difference between a time-consistent and a time-inconsistent solution (as in Example 2). Using the promised-utility approach one has to effectively recover $\overset{*}{\mu}_t$ – using the Envelope Theorem – to show these distortions. Furthermore, our co-state variable $\overset{*}{\mu}_t$ can also provide explicit interesting interpretations: in Example 1 the evolution of the μ's over time can be interpreted as time-varying Pareto weights; in Example 2 the behavior of the μ's is associated with a time-varying deadweight loss of taxation. We now turn to differences that appear in solving solutions to $\mathbf{PP}_\mu$.

At the core of our approach, there is a saddle-point functional equation, while the promised-utility approach works from a standard maximization Bellman-type equation. The continuation problem for us (namely $\mathbf{PP}_{\mu_1^*}$) is obtained by appropriately changing the weights μ in the objective function so that it is easy to guarantee the existence of a solution for any μ. The continuation problem for the promised-utility approach fixes one of the variables that enters the corresponding 'promise-keeping' constraint. This in effect changes the feasible set in period t and, as is well known, it opens up the possibility that the feasible set is empty and the continuation problem ill-defined. This is why under the promised-utility approach first of all one needs to compute the set of feasible promised utilities that render the continuation problem well-defined, and this feasible set has to be imposed on the Bellman equation. The computation of this feasible set can become daunting as problems become more complex.

Another difference is that our co-state variables have a well known initial condition $\mu_0 = \mu$, while the initial condition for the promised-utility approach needs to be solved for. We discuss these issues more concretely in the light of Example 2.

The key insight of the promised-utility approach is that a forward-looking constraint is converted into a standard backward-looking constraint. We can examine how this is done in Example 2. For ease of exposition we assume the exogenous shock g_t is i.i.d. and it can take ν possible values $\overline{g}^\kappa$, each with probability π^κ, for $\kappa = 1, ..., \nu$. Rewrite constraint (13) as

$$b_{t+1}\beta \sum_{\kappa=1}^{\nu} u'(c_{t+1}(\overline{g}^\kappa))\pi^\kappa = u'(c_t)(b_t - c_t) - e_t v'(e_t). \tag{18}$$

Equation (18) is the 'promise-keeping' constraint for consumption c_t (i.e. c_t is the co-state; alternatively, $u'(c_t)$ can be the co-state) and $c_{t+1}(\overline{g}^\kappa)$ is the promised consumption in period $t+1$ if $g_{t+1} = \overline{g}^\kappa$ is realized. By including all promised consumptions $(c_{t+1}(\overline{g}^1), ..., c_{t+1}(\overline{g}^\nu))$ in the vector of decision variables a_t, equation (18) becomes a special case of a standard (backward-looking) constraint (2). This suggests we can apply the Bellman equation to conclude that the problem is recursive as long as realized consumption $c_t(g_t)$ is included as a co-state variable.

Applying the Bellman equation to this reformulated problem is not straightforward. The Bellman equation may induce the planner to choose a $c_t(g_t)$ that cannot be supported by any taxation scheme that satisfies the budget constraint of the government from t onwards, so in this case the Bellman equation does not provide a feasible solution. To avoid this problem, one needs to compute the correspondence $\mathcal{C}_\kappa : R \to \mathcal{S}$, where $\mathcal{S}$ is a collection of subsets of R_+ such that if $c_{t+1}(\overline{g}^\kappa) \in \mathcal{C}_\kappa(b_t)$ and if $g_{t+1} = \overline{g}^\kappa$ then a continuation equilibrium tax $\{\tau_{t+j}\}_{j=1}^{\infty}$ exists for which a competitive equilibrium exists when $c_{t+1} = c_{t+1}(\overline{g}^\kappa)$ and given inherited government debt b_t. The correspondence $\mathcal{C}_\kappa(\cdot)$ is an endogenous object that needs to be computed before the Bellman equation is solved. This task becomes very complicated in higher dimensional problems. For example, if there were J types of consumers in the above Ramsey model, J promised consumptions would have to be carried over as state variables and in that case we would need to compute multidimensional sets $\mathcal{C}_\kappa(b) \subset R_+^J$. Even though considerable progress has been made in the computation of the correspondence $\mathcal{C}_\kappa$, either by improving algorithms or by redefining the problem at hand[15], this computation often leads to serious numerical difficulties.

As we have seen, the issue of computing a feasible set for promised consumption is entirely sidestepped in our approach. This is because any γ_{t-1}^{*1} gives a well-defined continuous objective function of $\mathbf{PP}_{\mu_t}$ in (16), so that this continuation problem always has a solution.

An additional advantage of the Lagrangian approach is that it leads to a reduction in the number of decision and state variables. We have only two decision variables (c_t, b_t) in Example 2 under our approach, while in the promised utility approach there are ν+1 decision variables $(c_{t+1}(\overline{g}^1), ..., c_{t+1}(\overline{g}^\nu), b_t)$.

For those with some experience in computing non-linear dynamic models, it is clear that the highest computational savings come from a reduction in the dimension of the state vector. In the above example, co-states are one-dimensional both in our approach (γ_{t-1}^1) and in the promised-utility

[15] See, for example, Abraham and Pavoni (2005) or Judd, Yeltekin and Conklin (2003).

approach (c_t). However, in some cases the recursive Lagrangian has many fewer state variables. Consider generalizing Example 2 to the case where the government issues one long bond that matures in M periods, as in Faraglia, Marcet and Scott (2014). In this case, the bond price depends on the expectation of marginal utility M periods ahead, so that when rewriting the budget constraint to convert it into a backward-looking constraint – as we did to derive (18) – we would now get

$$b_{t+1}^M \beta^M \sum_{\kappa=1}^{\nu^M} u'(c_{t+M}(\widetilde{g}^\kappa))\widetilde{\pi}^\kappa = u'(c_t)(b_{t-M+1}^M - c_t) - e_t v'(e_t), \tag{19}$$

where each $\widetilde{g}^\kappa$ is a *sequence* of possible realizations of $(g_{t+1}, ..., g_{t+M})$ and $\widetilde{\pi}^\kappa$ the probability of each sequence. Clearly, the co-state is now a vector of ν^M promised consumptions. For a 10-year bond, a quarterly model, even if g only takes two possible values so $\nu = 2$, the model has more than one trillion state variables. By comparison, the Lagrangian approach can be implemented with $2M+1 = 81$ state variables $(\gamma_{t-1}, ..., \gamma_{t-M}, b_t^M, ..., b_{t-M+1}^M, g_t)$[16]. Although there are ways of dealing with the above issues[17] a small set of state variables can be achieved naturally with our approach.

An additional difference is that the initial conditions for the co-state variables in our approach are known from the outset to be $\mu_0^0 = 1$, $\mu_0^1 = 0$, but in the promised-utility approach the initial condition is c_0, which needs to be solved for separately since it is an endogenous variable. It is well known that to find c_0 the Pareto frontier has to be downward sloping; otherwise the computations can become very cumbersome[18].

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

This section contains the main results of this paper, namely, that the maximization problem $\mathbf{PP}_\mu$ is equivalent to the **SPFE**, under fairly general conditions. As an intermediate step we use an infinite-dimensional one-period saddle-point problem $\mathbf{SPP}_\mu$. Proofs are deferred to the Appendix.

An outline of the results is as follows. We first show there is an equivalence between $\mathbf{PP}_\mu$ and $\mathbf{SPP}_\mu$. Theorem 1 shows that under standard concavity-convexity assumptions solutions of $\mathbf{PP}_\mu$ are solutions to the saddle-point problem $\mathbf{SPP}_\mu$ (we write this in shorthand as $\mathbf{PP}_\mu \Rightarrow \mathbf{SPP}_\mu$). Theorem 2 shows the converse, $\mathbf{SPP}_\mu \Rightarrow \mathbf{PP}_\mu$, assuming that $\mathbf{SPP}_\mu$ has a solution. Then we show that solutions to $\mathbf{PP}_{\mu_1^*}$ deliver the continuation of the solution to the original problem $\mathbf{PP}_\mu$ (Proposition 1).

Next, we show an equivalence between $\mathbf{SPP}_\mu$ and the **SPFE.** We show that if $\mathbf{SPP}_\mu$ has a solution for any (x, μ, s) then its value function and solution in period $t = 0$ satisfy the **SPFE** (in shorthand we write $\mathbf{SPP}_\mu \Rightarrow$**SPFE**; see Theorem 3). We finally show the converse, **SPFE**$\Rightarrow \mathbf{SPP}_\mu$.

[16]See Faraglia, Marcet and Scott (2014) for details.

[17]For example, Lustig et al (2008) provide a recursive formulation with long bonds by adding the yield curve as a state variable. The law of motion for the yield curve involves a long list of *forward-looking constraints*. The issue then becomes one of formulating a very high-dimensional feasible set for the yield curve which ensures that the continuation problem is well-defined.

[18]In Example 1, one may be interested in finding a 'fair' efficient allocation *ex-ante*. While this is trivial with our approach (just give the same initial weights in the **PP** problem), it becomes very tricky with the promised-utility approach, even with two agents, since the 'right promise' must be made to determine the initial conditions.

This last sufficiency result requires an *intertemporal consistency condition*, which is satisfied if the value function W of the **SPFE** is differentiable in μ, as is the case when the sequential saddle-point problem in the **SPFE** has a unique solution. Furthermore, if solutions are not unique (i.e. *saddle-point policy correspondence*) there is a sequential selection satisfying the *intertemporal consistency condition*, which can be verified generating a solution path from **SPFE** (Theorem 4 and its Corollary).

The four theorems together provide our central equivalence result: $\mathbf{PP}_\mu \iff$ **SPFE.** Each step of this equivalence obtains under different combinations of assumptions that we list in detail below. Only $\mathbf{PP}_\mu \Rightarrow \mathbf{SPP}_\mu$ (Theorem 1) relies on concavity-convexity assumptions, and only $\mathbf{SPFE} \Rightarrow \mathbf{SPP}_\mu$ (Corollary to Theorem 4) relies on a uniqueness assumption or, more generally, on a differentiability of the value function assumption as a sufficiency condition for our *intertemporal consistency condition*. The result $\mathbf{SPP}_\mu \Rightarrow \mathbf{PP}_\mu$ assumes the existence of a saddle-point. In Section 5 we show that under standard conditions there is a solution to the **SPFE** and, therefore, to $\mathbf{SPP}_\mu$ when $\mathbf{SPFE} \Rightarrow \mathbf{SPP}_\mu$ (Proposition 2). Theorems 1 and 2 are adaptations of optimization theory for infinite-dimensional optimization problems, while Proposition 1 and Theorems 3 and 4 are specific to our approach.

4.1 Assumptions on $\mathbf{PP}_\mu$

We consider the following set of assumptions:

A1. s_t takes values from 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 $p : X \times A \times S \to R^q$ and $\ell : X \times A \times S \to X$ are $\mathcal{S}$-measurable and, for any $s \in S$, they are continuous on (x, a). (**c**) For all (x, s), there is a program $\{\overline{a}_t\}_{t=0}^{\infty}$, with initial conditions (x, s), which satisfies constraints (2) and (3) for all $t \geq 0$.

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

A4. The functions $h_i^j(\cdot, \cdot, s)$, $i = 0, 1$, $j = 0, ..., l$, are $\mathcal{S}$-measurable and uniformly bounded and, for any $s \in S$, they are continuous on (x, a). Furthermore, $\beta \in (0, 1)$.

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: (**i**) satisfies constraints (2) with strict inequality for all $t \geq 0$, and (**ii**) satisfies constraints (3) with strict inequality for all $t \geq 0$.

Assumptions **A1-A4,** are standard, they hold in most applications, and we treat them as our basic assumptions. Note that **A3** and **A5** allow for sustained growth, although we maintain the assumption of bounded returns **A4**[19]. Assumptions **A5-A7** are not satisfied in some models of interest – for example, h_0^1 is not concave in Example 2. Assuming linearity of ℓ is without loss of generality, it allows for a reduction of the dimension of the state space[20]. However, these assumptions are only used in some of the results below. For example, the concavity assumptions **A5-A6** are not needed for sufficiency results, and assumption **A7** is a standard interiority assumption (the Slater condition), only needed to guarantee the existence of Lagrange multipliers and, therefore, that solutions to $\mathbf{PP}_\mu$ are the max component of saddle-points of the Lagrangean[21].

4.2 The intermediate step $\mathbf{SPP}_\mu$ and its relationship with $\mathbf{PP}_\mu$

We now build a Lagrangian that gives the solution to $\mathbf{PP}_\mu$. In Section 2, by adding linear combinations of forward-looking constraints for *all* periods to the objective function of $\mathbf{PP}_\mu$, we defined $\mathcal{L}_\mu$ as the objective function of an infinite-horizon saddle-point. While this construction provided intution, it is convenient for the formal analysis to proceed differently and to only add the forward-looking constraint for the period $t = 0$, with its multiplier γ to the objective function of $\mathbf{PP}_\mu$, leaving the remaining forward-looking constraints for $t \geq 1$ in the definition of the feasible set. Proceeding in this way and after some simple algebra, the resulting *saddle-point problem* can be written as

$$\mathbf{SPP}_\mu : \quad SV(x,\mu,s) = \underset{\inf_{\gamma \in R_+^{l+1}}, \sup_{\{a_t\}_{t=0}^\infty}}{\mathrm{SP}} \Big\{ \mu h_0(x_0,\, a_0,\, s_0) + \gamma h_1(x_0,\, a_0,\, s_0)$$

$$+\beta\ \mathrm{E}_0 \sum_{j=0}^{l} \varphi^j(\mu,\, \gamma) \sum_{t=0}^{N_j} \beta^t\ h_0^j(x_{t+1},\, a_{t+1},\, s_{t+1}) \Big\} \tag{20}$$

$$\text{s.t. } x_{t+1} = \ell(x_t,\, a_t,\, s_{t+1}),\ \ p(x_t, a_t,\, s_t) \geq 0, \quad t \geq 0, \tag{21}$$

$$\mathrm{E}_t \sum_{n=1}^{N_j+1} \beta^n h_0^j(x_{t+n},\, a_{t+n},\, s_{t+n}) + h_1^j(x_t,\, a_t,\, s_t) \geq 0,\ j = 0, ..., l,\ t \geq 1, \tag{22}$$

given $x_0 = x$, $s_0 = s$. In particular, the path $(\gamma^*, \{a_t^*\}_{t=0}^\infty)$ solves $\mathbf{SPP}_\mu$ at (x, s) if $\gamma^* \in R_+^{l+1}$, $\{a_t^*\}_{t=0}^\infty$ satisfies (21) - (22), and for any $\gamma \in R_+^{l+1}$ and $\{a_t\}_{t=0}^\infty$ satisfying (21) - (22):

[19] Our theory can be extended to unbounded returns in the same way that standard dynamic programming can (see, for example, Stokey et al (1989) 4.3 - 4.4). For simplicity, we focus here on the case of bounded returns.

[20] In the case that a state variable at time t is a non-linear function of past state variables, one can always embed this non-linearity in the decision variables x_t and then select this state variable for the next period through ℓ.

[21] $\mathbf{PP}_\mu$ is an infinite-dimensional maximization problem. One can show that for any (x, s) there exists a solution if **A1-A6** are satisfied (Proposition 1 in Marcet and Marimon 2011).

$$\mu h_0(x_0,\, a_0^*,\, s_0) + \gamma h_1(x_0,\, a_0^*,\, s_0) + \beta \ \mathrm{E}_0 \sum_{j=0}^{l} \varphi^j(\mu,\, \gamma) \sum_{t=0}^{N_j} \beta^t \ h_0^j(x_{t+1}^*,\, a_{t+1}^*,\, s_{t+1})$$

$$\geq \ \mu h_0(x_0,\, a_0^*,\, s_0) + \gamma^* h_1(x_0,\, a_0^*,\, s_0) + \beta \ \mathrm{E}_0 \sum_{j=0}^{l} \varphi^j(\mu,\, \gamma^*) \sum_{t=0}^{N_j} \beta^t \ h_0^j(x_{t+1}^*,\, a_{t+1}^*,\, s_{t+1}) \tag{23}$$

$$\geq \ \mu h_0(x_0,\, a_0,\, s_0) + \gamma^* h_1(x_0,\, a_0,\, s_0) + \beta \ \mathrm{E}_0 \sum_{j=0}^{l} \varphi^j(\mu,\, \gamma^*) \sum_{t=0}^{N_j} \beta^t \ h_0^j(x_{t+1},\, a_{t+1},\, s_{t+1}). \tag{24}$$

That is, solutions to $\mathbf{SPP}_\mu$ at (x, s) are *saddle-points* in period zero. The following result says that a solution to the maximum problem is also a solution to $\mathbf{SPP}_\mu$. This result follows from the standard theory of constrained optimization in linear vector spaces. As in the standard theory, convexity and concavity assumptions (**A5** and **A6**), as well as an interiority assumption (**A7**), are necessary to obtain the result.

Theorem 1 ($\mathbf{PP}_\mu \Rightarrow \mathbf{SPP}_\mu$). Assume **A1-A6** and **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 $\mathbf{PP}_\mu$; i.e. $SV(x, \mu, s) = V_\mu(x, s)$.

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

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

Theorem 2 ($\mathbf{SPP}_\mu \Rightarrow \mathbf{PP}_\mu$). Assume **A1.** Assume $(\boldsymbol{a}^*,\, \gamma^*)$ is a solution to $\mathbf{SPP}_\mu$ for $(x,\, \mu,\, s) \in X \times R_+^{l+1} \times S$. Then $\boldsymbol{a}^*$ is a solution to $\mathbf{PP}_\mu$ for initial conditions $(x,\, s)$. Furthermore, $V_\mu(x, s) = SV(x, \mu, s)$.

Proof: See Appendix A.

Note that Theorem 2 is a sufficiency theorem 'almost free of assumptions.' Once the existence of a solution to $\mathbf{SPP}_\mu$ is granted, assumptions **A2** to **A7** are not needed. In particular, while concavity and interiority assumptions are needed to prove necessity (Theorem 1), they are not needed to prove sufficiency (Theorem 2). However, without interiority, or concavity, there may not be a saddle-point. Theorem 2 plays two roles in our approach: first, as an intermediate step connecting the $\mathbf{PP}_\mu$ and the **SPFE** solutions; second, it allows us to show that $\mathbf{PP}_{\mu_1^*}$ is the continuation problem of $\mathbf{PP}_\mu$.

As we have discussed in Section 2, the solution to $\mathbf{PP}_\mu$ is often time-inconsistent. If $\{a_t^*\}_{t=0}^\infty$ solves $\mathbf{PP}_\mu$, with initial conditions (x, s), the continuation of the optimal solution $\{a_t^*\}_{t=1}^\infty$ does not solve $\mathbf{PP}_\mu$ for initial conditions (x_1^*, s_1) where $x_1^* = \ell(x,\, a_0^*,\, s_1)$ if the *forward-looking* constraints are binding (i.e. $\gamma_0^* \neq 0$). The following proposition says that the continuation $\{a_t^*\}_{t=1}^\infty$ actually solves a problem with a different objective function with adjusted weights, namely $\mathbf{PP}_{\mu_1^*}$ with initial conditions (x_1^*, s_1), where $\mu_1^* = \varphi(\mu,\, \gamma_0^*)$.

Proposition 1 (Continuation Problem): Assume **A1** and **A2 (c).** Assume $(\boldsymbol{a}^*, \gamma^*)$ is a solution to $\mathbf{SPP}_\mu$ for $(x, s) \in X \times S$. The continuation $\{a_t^*\}_{t=1}^\infty$ solves $\mathbf{PP}_{\mu_1^*}$ with initial conditions (x_1^*, s_1) a.s. in s_1, where $x_1^* = \ell(x, a_0^*, s_1)$.

Proof: See Appendix A.

4.3 The relationship between $\mathbf{SPP}_\mu$ and the SPFE:

Necessity

We now establish the relationship between $\mathbf{SPP}_\mu$ and the **SPFE.** Given a value function W satisfying the **SPFE** (8) in *any possible state* $(x, \mu, s) \in X \times R_+^{l+1} \times S$, the corresponding *saddle-point policy correspondence* (*SP policy correspondence*) $\Psi : X \times R_+^{l+1} \times S \to A \times R_+^{l+1}$ is:

$$\begin{aligned}
\Psi_W(x,\, \mu,\, s) \quad = \quad & \{(a^*, \gamma^*) \in X \times R_+^{l+1} \text{ satisfying } p(x, a^*,\, s) \geq 0 \text{ s.t.} \\
& \mu h_0(x,\, a^*,\, s) + \gamma h_1(x,\, a^*,\, s) + \beta \mathrm{E}\left[W(\ell(x,\, a^*,\, s'),\, \varphi(\mu,\, \gamma),\, s') |\, s\right] \\
& \geq \mu h_0(x,\, a^*,\, s) + \gamma^* h_1(x,\, a^*,\, s) + \beta \mathrm{E}\left[W(\ell(x,\, a^*,\, s'),\, \varphi(\mu,\, \gamma^*),\, s') |\, s\right] \quad (25) \\
& \geq \mu h_0(x,\, a,\, s) + \gamma^* h_1(x,\, a,\, s) + \beta\, \mathrm{E}\left[W(\ell(x,\, a,\, s'),\, \varphi(\mu,\, \gamma^*),\, s') |\, s\right] \quad (26) \\
& \text{for all} \quad (a, \gamma) \in X \times R_+^{l+1} \text{ satisfying } p(x, a,\, s) \geq 0\}.
\end{aligned}$$

If $\Psi_W(x,\, \mu,\, s) \neq \varnothing$, then there exists a *saddle-point* at $(x,\, \mu,\, s)$ and $W(x,\, \mu,\, s)$ is well-defined. If, in addition, Ψ_W is single valued, we denote it by ψ_W (i.e. $(a^*, \lambda^*) = \psi_W(x,\, \mu,\, s)$)[22], and we call it a *saddle-point policy function* (*SP policy function*).

The following theorem says that SV satisfies **SPFE**:

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

Proof: See Appendix C.

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 $\mathbf{SPP}_\mu$ at all possible (x, μ, s) is assumed.

Sufficiency

We now turn to our sufficiency theorem: **SPFE** $\Rightarrow \mathbf{SPP}_\mu$. We do not assume that the value function, W, satisfying **SPFE** is differentiable but only that it is continuous in (x, μ) and convex and homogeneous of degree one in μ, for every s. To see that this is the natural class of value functions to consider, note that, by Theorems 1 and 2, if $\boldsymbol{a}^*$ is a solution to $\mathbf{PP}_\mu$ at (x, s), and $(x_0, s_0) = (x, s)$, then

$$SV(x, \mu, s) = V_\mu(x, s) = \mathrm{E}_0 \sum_{j=0}^{l} \sum_{t=0}^{N_j} \beta^t \mu^j h_0^j(x_t^*,\, a_t^*,\, s_t).$$

[22]We often simplify notation by writing a^* for $a^*(x,\, \mu,\, s)$.

It can easily be established that, for any s, SV is continuous in (x, μ) and convex and homogeneous of degree one in μ, and if **A6** is assumed SV is also concave in x (see Appendix B). In particular, SV has the following representation:

$$SV(x, \mu, s) = \sum_{j=0}^{l} \mu^j sv^j(x, \mu, s),$$

where $sv^j(x, \mu, s) \equiv \mathrm{E}_0 \sum_{t=0}^{N_j} \beta^t h_0^j(x_t^*,\, a_t^*,\, s_t)$. We call this the *Euler representation*; the corresponding Euler's Theorem assumes that SV is differentiable in μ and, therefore, sv^j is the partial derivative with respect to μ^j. However, this result generalizes to the non-differentiable case; that is, given a function W, continuous in (x, μ) and convex and homogeneous of degree one in μ, $\partial_\mu W(x,\, \mu,\, s)$ denotes the *subdifferential* of W at $(x,\, \mu,\, s)$ with respect to μ – i.e.

$$\partial_\mu W(x,\, \mu,\, s) = \left\{ \omega \in R^{l+1} \mid W(x,\, \widetilde{\mu},\, s) \geq W(x,\, \mu,\, s) + (\widetilde{\mu} - \mu)\, \omega \text{ for all } \widetilde{\mu} \in R_+^{l+1} \right\}.$$

As we show in Lemma 1(i), W also has a *Euler representation*: $W(x,\, \mu,\, s) = \mu\omega(x,\, \mu,\, s)$, for any $\omega(x,\, \mu,\, s) \in \partial_\mu W(x,\, \mu,\, s)$. In our case, this representation will allow us to write the first-order (*Kuhn-Tucker*) conditions of (25) as

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

for $j = 0, ..., l$ and $\omega^j(\ell(x,\, a,\, s'),\, \varphi(\mu,\, \gamma^*),\, s') \in \partial_\mu W(\ell(x,\, a,\, s'),\, \varphi(\mu,\, \gamma^*),\, s')$ (see Lemma 1(ii)). We call these ω^j *supporting selections* when the corresponding Euler representation of W satisfies the *saddle-point* inequalities (25) and (26). We can now state our main sufficiency theorem:

Theorem 4. Assume W, satisfying the **SPFE**, is continuous in (x, μ) and convex and homogeneous of degree one in μ, for every s. Let Ψ_W be *the SP policy correspondence* associated with W which generates a solution $(\mathbf{a}^*, \boldsymbol{\gamma}^*)_{(x,\mu,s)}$ satisfying $\lim_{t\to\infty} \beta^t W(x_t^*,\, \mu_t^*,\, s_t) = 0$ and supported by selections $\omega(x_t^*,\, \mu_t^*, s_t) \in \partial_\mu W(x_t^*,\, \mu_t^*,\, s_t)$. If, for $t \geq 0$ and $j = 0, ..., k$, the following *intertemporal consistency condition* is satisfied

$$\omega^j(x_t^*,\, \mu_t^*,\, s_t) = h_0^j(x^*,\, a_t^*,\, s) + \beta\ \mathrm{E}\left[\omega^j(x_{t+1}^*,\, \mu_{t+1}^*,\, s_{t+1})|\, s_t\right], \tag{27}$$

$$\text{and} \lim_{t\to\infty} \beta^t \omega(x_t^*,\, \mu_t^*,\, s_t) = 0, \tag{28}$$

then $(\mathbf{a}^*, \gamma_0^*)_{(x,\mu,s)}$ is also a solution to $\mathbf{SPP}_\mu$ and $\mathbf{PP}_\mu$ at (x, s), and $V_\mu(x,\, s) = SV(x,\, \mu,\, s) = W(x,\, \mu,\, s)$.

Proof: See Appendix C.

Condition (28) is fairly innocuous since for every $(x_t^*\, \mu_t^*, s_t)$ the subdifferential, $\partial_\mu W(x_t^*,\, \mu_t^*,\, s_t)$ is bounded, which in most applications implies (28) – e.g. if all ω^j, $j = 0, ..., k$, have the same sign. In contrast, the *intertemporal consistency condition* (27) is an important necessary condition. As has already been discussed, the 'promised utility' approach imposes (27) as a constraint, which can be an inconvenient restriction to impose in practice. Our sufficiency results show that this inconvenience is sidestepped when using our approach. In particular, our consistency condition (27) only needs to be checked *ex-post*, not imposed *ex-ante*, which makes a difference computationally. The following corollary to Theorem 4 shows that given a solution to the **SPFE** it is always possible to find supporting selections satisfying the *intertemporal consistency condition* (27). The corollary also provides sufficiency conditions for (27) to be satisfied:

Corollary (SPFE $\Longrightarrow$ SPP$_\mu$): Assume W satisfies **SPFE** and the assumptions of Theorem 4. Let Ψ_W be *the SP policy correspondence* associated with W which generates a solution $(\mathbf{a}^*, \boldsymbol{\gamma}^*)_{(x,\mu,s)}$ satisfying $\lim_{t\to\infty} \beta^t W(x_t^*,\, \mu_t^*,\, s_t) = 0$: *i)* there are supporting selections $\omega^*(x_t^*,\, \mu_t^*, s_t) \in \partial_\mu W(x_t^*,\, \mu_t^*,\, s_t)$ satisfying the *intertemporal consistency condition* (27); *ii)* if, at any $(x_t^*,\, \mu_t^*,\, s_t)$, W is differentiable in μ then (27) is satisfied; *iii)* if $(\mathbf{a}^*)_{(x,\mu,s)}$ is uniquely determined and its supporting selections satisfy (28), then $(\mathbf{a}^*, \boldsymbol{\gamma}^*)_{(x,\mu,s)}$ is also a solution to **SPP**$_\mu$ at (x, s).

Proof: See Appendix C.

Note that, if W is not differentiable in μ (27) may not be satisfied by a solution to **SPFE**[23]. The proofs of Theorem 4 and of its Corollary make use of the following lemma which provides, in (i), the *Euler representation of* W when it is not differentiable in μ and, in (ii), a convenient *Kuhn-Tucker* characterization of the **SPFE** *saddle-point* conditions.

Lemma 1. Let W be continuous in (x, μ) and convex and homogeneous of degree one in μ, for every s.

i) If $W(x,\, \mu,\, s)$ is finite, $\partial_\mu W(x,\, \mu,\, s) \neq \varnothing$ and if $\omega(x,\, \mu,\, s) \in \partial_\mu W(x,\, \mu,\, s)$ then $W(x,\, \mu,\, s) = \mu\omega(x,\, \mu,\, s)$ and, for all $\lambda > 0$, $\omega(x,\, \mu,\, s) \in \partial_\mu W(x,\, \lambda\mu,\, s)$. Furthermore, W is differentiable in μ at $(x,\, \mu,\, s)$ if, and only if, $\partial_\mu W(x,\, \mu,\, s)$ is a singleton.

ii) $(a^*, \gamma^*) \in \Psi_W(x,\, \mu,\, s)$ *if and only if*, for all s' reached from s, there is a $\omega(x^{*\prime},\, \mu^{*\prime}, s') \in \partial_\mu W(x^{*\prime},\, \mu^{*\prime}, s')$ with $x^{*\prime} = \ell(x,\; a^*,\; s')$ and $\mu^{*\prime} = \varphi(\mu,\, \gamma^*)$, such that:

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

$$\geq \mu h_0(x,\, a,\, s) + \gamma^* h_1(x,\, a,\, s) + \beta \,\mathrm{E}\left[\varphi(\mu,\, \gamma^*)\omega(x',\, \varphi(\mu,\, \gamma^*),\, s')|\, s\right], \tag{29}$$

for all $a \in A$ and $x' = \ell(x,\, a,\, s')$ satisfying $p(x, a,\, s) \geq 0$, and, for $j = 0, ..., l$,

$$h_1^j(x,\, a^*,\, s) + \beta \,\mathrm{E}\left[\omega^j(x^{*\prime},\, \varphi(\mu,\, \gamma^*),\, s')|\, s\right] \;\; \geq \;\; 0, \tag{30}$$

$$\gamma^{*j}\left[h_1^j(x,\, a^*,\, s) + \beta \,\mathrm{E}\left[\omega^j(x^{*\prime},\, \varphi(\mu,\, \gamma^*),\, s')|\, s\right]\right] \;\; = \;\; 0. \tag{31}$$

[23] Marimon and Werner (2017) show that this problem is just a manifestation of a more general one: in standard dynamic programming, Lagrange multipliers associated with binding constraints may not be unique and, correspondingly, the value function may not be differentiable. In this case, the intertemporal *Euler equations* may not be satisfied. They show that an *envelope condition* – which amounts to making 'consistent selections' from the *envelope subdifferential* – is necessary and sufficient to recover 'the necessity' of the Euler equations. In our approach, (27) is the *Euler equation* with respect to $\mu^j, j = 0, \ldots, k$. Their *envelope condition* refers to the Lagrange multipliers associated with the constraints $\mu_{t+1}^j \geq \mu_t^j$ and it implies our *intertemporal consistency condition* (27). They also show how to extend our **SPFE** to the case where W is not differentiable in μ, by further extending the co-state with these Lagrange multipliers.

Furthermore, (30) and (31) are satisfied if and only if the following inequality is satisfied:

$$\begin{aligned} & \mu h_0(x,\, a^*,\, s) + \gamma h_1(x,\, a^*,\, s) + \beta \, \mathrm{E}\left[\varphi(\mu,\, \gamma)\omega(x',\, \varphi(\mu,\, \gamma^*),\, s')|\, s\right] \\ \geq \;\; & \mu h_0(x,\, a^*,\, s) + \gamma^* h_1(x,\, a^*,\, s) + \beta \, \mathrm{E}\left[\varphi(\mu,\, \gamma^*)\omega(x',\, \varphi(\mu,\, \gamma^*),\, s')|\, s\right] \end{aligned} \tag{32}$$

for all $\gamma \in R_+^{l+1}$.

Proof: See Appendix B, which also includes a more detailed discussion of properties of continuous, convex and homogeneous functions.

Note that Lemma 1 (*ii*) says that both (25) and (32) share the same *Kuhn-Tucker* conditions (30) and (31)[24]. Lemma 1 (*i*) and (31) imply that, if $(a_t^*, \gamma_t^*) \in \Psi_W(x_t^*,\, \mu_t^*,\, s_t)$, we can write

$$W(x_t^*,\, \mu_t^*,\, s_t) = \mu_t^* h_0(x_t^*,\, a_t^*,\, s_t) + \gamma_t^* h_1(x_t^*,\, a_t^*,\, s_t) + \beta \, \mathrm{E}\left[W(x_{t+1}^*,\, \varphi(\mu_t^*,\, \gamma_t^*),\, s_{t+1})|\, s_t\right] \tag{33}$$

as[25]

$$\mu_t^* \omega_t(x_t^*,\, \mu_t^*,\, s_t) = \mu_t^* h_0(x_t^*,\, a_t^*,\, s_t) + \beta \sum_{j=0}^{k} \mu_t^{*j} \mathrm{E}\left[\omega_t^j(x_{t+1}^*,\, \varphi(\mu_t^*,\, \gamma_t^*),\, s_{t+1})|\, s_t\right], \tag{34}$$

where we have used the fact that, by (31), we can eliminate the binding *forward-looking* constraints – more precisely:

$$\begin{aligned} W(x_t^*,\, \mu_t^*,\, s_t) &= \mu_t^* \omega_t(x_t^*,\, \mu_t^*,\, s_t) \\ &= \mu_t^* h_0(x_t^*,\, a_t^*,\, s_t) + \gamma_t^* h_1(x_t^*,\, a_t^*,\, s_t) + \beta \, \mathrm{E}\left[\varphi(\mu_t^*,\, \gamma_t^*)\omega_t(x_{t+1}^*,\, \varphi(\mu_t^*,\, \gamma_t^*),\, s_{t+1})|\, s_t\right] \\ &= \mu_t^* h_0(x_t^*,\, a_t^*,\, s_t) + \beta \sum_{j=0}^{k} \mu_t^{*j} \mathrm{E}\left[\omega_t^j(x_{t+1}^*,\, \varphi(\mu_t^*,\, \gamma_t^*),\, s_{t+1})|\, s_t\right] \\ &\quad + \gamma_t^* \left[h_1(x_t^*,\, a_t^*,\, s_t) + \beta \, \mathrm{E}\omega_t(x_{t+1}^*,\, \varphi(\mu_t^*,\, \gamma_t^*),\, s_{t+1})|\, s_t\right] \\ &= \mu_t^* h_0(x_t^*,\, a_t^*,\, s_t) + \beta \sum_{j=0}^{k} \mu_t^{*j} \mathrm{E}\left[\omega_t^j(x_{t+1}^*,\, \varphi(\mu_t^*,\, \gamma_t^*),\, s_{t+1})|\, s_t\right]. \end{aligned}$$

Finally, note that (27) being satisfied means that $\omega_t(x_{t+1}^*,\, \mu_{t+1}^*,\, s_{t+1}) = \omega_{t+1}(x_{t+1}^*,\, \mu_{t+1}^*,\, s_{t+1}) = \omega(x_{t+1}^*,\, \mu_{t+1}^*,\, s_{t+1})$ for all $(t+1, s_{t+1})$.

5 Existence of saddle-point value functions

In this section, we address the issue of the existence of value functions satisfying the **SPFE** (Proposition 2 (*i*)). The existence of saddle-points is needed to show that there is a well-defined contraction

[24] Note the subtle difference between (25) and (32): in the latter γ^* is an argument of ω in both sides of the inequality, making the function to be minimized linear in γ, as with standard Lagrange multipliers.

[25] The time subindex in ω denotes the timing of the selection. That is, $\omega_t^j(x_{t+1}^*,\, \varphi(\mu_t^*,\, \gamma_t^*),\, s_{t+1})$ on the right-hand side of (34) denotes the selection from $\partial_\mu W(x_{t+1}^*,\, \varphi(\mu_t^*,\, \gamma_t^*),\, s_{t+1})$ when a_t^* is chosen, while $\omega_{t+1}^j(x_{t+1}^*,\, \varphi(\mu_t^*,\, \gamma_t^*),\, s_{t+1})$ is the selection corresponding to the left-hand side of (34) the following period. When W is not differentiable at $(x_{t+1}^*,\, \varphi(\mu_t^*,\, \gamma_t^*),\, s_{t+1})$ these two selections may not be the same, in which case the *intertemporal consistency condition* (27) would not be satisfied. Messner and Pavoni (2004) provide an example of such *inconsistency*.

mapping generalizing the *Contraction Mapping Theorem* to a *Dynamic Saddle-Point Problem* corresponding to the **SPFE** (Proposition 2 (*iii*))[26].

We first define the space of bounded value functions (in x) which are convex and homogeneous of degree one (in μ):

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

and we also define its subspace of concave functions (in x): $\mathcal{M}_{bc} = \{W \in \mathcal{M}_b$ and *iii*) $W(\cdot,\,\mu,\,s)$ is concave$\}$. 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 these are complete metric spaces, and therefore, suitable spaces for the *Contraction Mapping Theorem.* As we have seen in Section 4[27], $SV(x,\mu,s)$, the value of $\mathbf{SPP}_\mu$ with initial conditions (x,s), is an element of $\mathcal{M}_b$ whenever **A2 - A4** are satisfied (and $SV \in \mathcal{M}_{bc}$ if in addition **A5 - A6** are satisfied).

Let $\mathcal{M}$ denote either $\mathcal{M}_b$ or $\mathcal{M}_{bc}$. Then the **SPFE** defines a saddle-point operator $T^* : \mathcal{M} \longrightarrow \mathcal{M}$ given by

$$(T^*W)(x,\mu,s) = \underset{\min_{\gamma\geq 0},\max_a}{\text{SP}} \left\{\mu h_0(x,\,a,\,s) + \gamma h_1(x,\,a,\,s) + \beta\ \text{E}\left[W(x',\,\mu',\,s')|\,s\right]\right\} \tag{35}$$

$$\begin{aligned}&\text{s.t. } x' = \ell(x,\,a,\,s'),\ p(x,a,\,s) \geq 0, \\ &\text{and } \mu' = \varphi(\mu,\,\gamma).\end{aligned}$$

In defining T^* as a *saddle-point operator* we have implicitly assumed that there is a *saddle-point* (a^*,γ^*) satisfying:

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

$$\forall\ \gamma \in \mathcal{R}_+^{l+1},\ \mu' = \varphi(\mu,\,\gamma) \text{ and } a \text{ with } p(x,a,\,s) \geq 0,\ x' = \ell(x,\,a,\,s').$$

[26]As is well known, if the solutions to the *primal* (sup inf) and the *dual* (inf sup) problems have the same value, then there is a *saddle-point*, however this identity between *primal* and *dual* values is not necessary for the existence of a *saddle-point.* We follow a direct approach of proving existence of *saddle-points* without relying on duality arguments. Messner, Pavoni and Sleet (2013) study the alternative duality approach in an infinite-dimensional formulation of a – in some dimensions, generalized – version of our framework. It should also be noted that with infinite-dimensional state and co-state variables computational complexity may severely restrict the range of applications of our approach

[27]See also Lemma 1A in Appendix B.

To guarantee that the T^* operator preserves measurability we strengthen assumption **A1**:

A1b. s_t takes values from a compact and convex set $S \subset R^K$. $\{s_t\}_{t=0}^{\infty}$ is a Markovian stochastic process defined on the probability space $(S_\infty, \mathcal{S}, P)$ with transition function Q on $(S, \mathcal{S})$ satisfying the Feller property[28].

Given that W is convex and homogeneous of degree one, it has a *Euler representation*: $W(x',\, \mu',\, s') = \mu'\omega(x',\, \mu',\, s')$ (recall Lemma 1(i)). We can then consider the following *Dynamic Programming Problem* (**DPP**) at (x, μ, s), given γ and $\mu' = \varphi(\mu,\, \gamma)$:

$$\max_{a} \left\{ \mu h_0(x,\, a,\, s) + \beta \;\mathrm{E}\left[\sum_{j=0}^{k} \mu^j \omega^j(x',\, \mu',\, s') |\, s \right] \right\} \tag{36}$$

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

$$\text{and } h_1^j(x,\, a,\, s) + \beta \;\mathrm{E}\left[\omega^j(x',\, \mu',\, s')|\, s \right] \geq 0,\; j = 0, ..., l. \tag{37}$$

DPP is a standard constrained optimization problem, which, given our assumptions – and provided that an interiority condition associated with (37) is satisfied – has a solution $a^*(x, \mu, s; \gamma)$ and Lagrange multipliers $\gamma^*(x, \mu, s; \gamma)$ corresponding to (37). In particular, (a^*, γ^*) is a *saddle-point* of the operator T^* if it satisfies: $a^* = a^*(x, \mu, s; \gamma^*)$ and $\gamma^* = \gamma^*(x, \mu, s; \gamma^*)$. That is, we can decompose the problem of the existence of *saddle-point* into two problems: a constrained maximization problem (36) and a fixed-point condition that the Lagrange multipliers, entering the *forward-looking constraints,* must satisfy: $\gamma^* = \gamma^*(x, \mu, s; \gamma^*)$. We follow this approach, and the interiority condition for *forward-looking constraints* (37) that we assume is:

IC. W satisfies the *interiority condition* at $(x,\, \mu,\, s) \in X \times R_+^{l+1} \times S$ if for any $\gamma \in R_+^{l+1}, \gamma \neq 0$, there exists $\widetilde{a} \in A$, satisfying $p(x,\, \widetilde{a},\, s) > 0$, and, for all s' reached from s, there is a $\omega(\ell(x,\; \widetilde{a},\, s'),\, \varphi(\mu,\, \gamma), s') \in \partial_\mu W(\ell(x,\; \widetilde{a},\, s'),\, \varphi(\mu,\, \gamma), s')$, such that $\gamma\left[h_1(x, \widetilde{a}, s) + \beta \mathrm{E}\left[\omega(\ell(x,\; \widetilde{a},\, s'),\, \varphi(\mu,\, \gamma), s')|\, s\right]\right] > 0$.

Assuming **A1 - A6** and **A7**(i), **IC** guarantees that the Kuhn-Tucker conditions of Lemma 1, (30) and (31) – together with the corresponding Kuhn-Tucker conditions for the feasibility constraints (2) – are necessary and sufficient for $a^*(x, \mu, s; \gamma^*)$ to be a solution to (36) at $\gamma^* = \gamma^*(x, \mu, s; \gamma^*)$. **IC** is not a restrictive assumption[29] if the original **PP** has interior solutions, as we assume with **A7**(ii)[30]. It follows from **A7**(ii) in the case of intertemporal one-period constraints, where $\omega^j(\widetilde{x},\, \mu', s') = h_0^j(\widetilde{x},\, \mu', s')$. In

[28] Recall that Q satisfies the Feller property if whenever f is bounded and continuous on S, the function Tf given by $(Tf)(s) = \int f(s')Q(s, ds')$, for all $s \in S$ is also bounded and continuous on S. **A1** can be alternatively strengthened by assuming that S is countable and $\mathcal{S}$ is the σ-algebra containing all the subsets of S (see Stokey, et al. (1989) 9.2).

[29] Although it is a necessary condition for the existence of the separation hyperplane defining the saddle-point which in our context has a specific meaning. Without it, one can show – in a version of Example 1 – that the autarkic solution can satisfy the saddle-point conditions of SPFE, although SPFE is not well-defined and the resulting value function violates the transversality condition $\lim_{T \longrightarrow \infty} \beta^T W = 0$ (Charles Brendon, personal communication, March 2015).

[30] **IC** is a generalization to *forward-looking constraints* of *Karlin's condition*, which is equivalent to the standard *Slater's condition*: there exists a feasible solution $\tilde{a}$ satisfying (37) with inequality (see, for example, Takayama (1974)).

the case of intertemporal participation constraints – such as Example 1 – **IC** requires that at any (x, μ, s) there is positive surplus, according to W, to be shared and, therefore, it is efficient to continue the partnership[31]. We strengthen the interiority condition **IC** in order to uniformly bind the Kuhn-Tucker multipliers, which allows Kakutani's Fixed Point Theorem to be applied:

SIC. W satisfies the *strict interiority condition* at $(x,\, \mu,\, s) \in X \times R_{+}^{l+1} \times S$ if it satisfies **IC** and there exists a $\varepsilon > 0$, such that $\gamma \left[h_1(x, \widetilde{a}, s) + \beta \mathrm{E} \left[\omega(\ell(x,\ \widetilde{a},\ s'),\ \varphi(\mu,\ \gamma), s') \middle| \, s \right] \right] \geq \varepsilon \left\| \gamma \right\|.$

Condition **SIC** implies that there is a $C > 0$, such that, if γ^* is a Kuhn-Tucker multiplier for $(x, \mu, s; \gamma)$, then $\|\gamma^*\| \leq C \|\mu\|$, in other words if **SIC** is satisfied Kuhn-Tucker multipliers are uniformly bounded, relative to the co-state μ (see the proof of Proposition 2 in Appendix D). We can now state the main result of this section:

Proposition 2. Assume **A1b** and **A2-A5** and **A7**(i), and **A6** when $\mathcal{M}$ refers to $\mathcal{M}_{bc}$.

i) Let $W \in \mathcal{M}_{bc}$ and assume, in addition, **SIC.** For all $(x, \mu, s) \in X \times R_{+}^{l+1} \times S$, there exists $(\mathbf{a}^*, \boldsymbol{\gamma}^*)_{(x,\mu,s)}$ generated by $\Psi_W(x, \mu, s)$; i.e. $(\mathbf{a}^*, \boldsymbol{\gamma}^*)_{(x,\mu,s)}$ satisfies (29) - (31). Furthermore, if **A6s** is assumed, then $(\mathbf{a}^*)_{(x,\mu,s)}$ is uniquely determined.

ii) Let $W \in \mathcal{M}$ if, for all $(x, \mu, s) \in X \times R_{+}^{l+1} \times S$, $\Psi_W(x, \mu, s) \neq \varnothing$. Then $T^*W \in \mathcal{M}$, i.e. $T^* : \mathcal{M} \longrightarrow \mathcal{M}$.

iii) Let $W \in \mathcal{M}$, if, for all $(x, \mu, s) \in X \times R_{+}^{l+1} \times S$, $\Psi_W(x, \mu, s) \neq \varnothing$. Then $T^* : \mathcal{M} \longrightarrow \mathcal{M}$ is a contraction mapping of modulus β.

Proof: See Appendix D.

Proposition 2 (*i*) provides conditions for the existence of a *saddle-point*, the proof follows the approach described above of decomposing the existence problem into two; (*ii*) completes the proof that the **SPFE** mapping is well defined by showing that T^* maps $\mathcal{M}$ onto itself, and finally (*iii*) shows it is a contraction mapping. This last result (*iii*) follows from the second (*ii*), Feller's property (**A1b**), and the fact that T^* satisfies Blackwell's sufficiency conditions for a contraction.

Proposition 2 shows how the standard dynamic programming results on the existence and uniqueness of a value function and the corresponding existence of optimal solutions generalise to our saddle-point dynamic programming approach, provided that an interiority condition is satisfied (e.g. **SIC**). In standard dynamic programming, when solutions are not unique a selection must be chosen from the optimal correspondence. This is also true in our *saddle-point* formulation. However, in this case the selection must satisfy an *intertemporal consistency condition* in order to be a solution of the original $\mathbf{PP}_\mu$ problem[32]. Finally, Proposition 2 also shows that *the contraction property* – very practical for computing value functions – also extends to our *saddle-point Bellman equation operator*.

[31] Alternatively, one can extend our formulation to include terminal nodes where there is no more surplus to be shared and the continuation values are just the constrained participation values.

[32] Condition (27) in Theorem 4; see also Corollary to Theorem 4 and footnote (23).

6 Related work

Forward-looking constraints are pervasive in dynamic economic models and precedents of our approach can be found in Epple, Hansen and Roberds (1985), Sargent (1987) and Levine and Currie (1987), who introduced Lagrange multipliers as co-state variables in linear-quadratic Ramsey problems. Similarly, studies of optimal monetary policy in sticky price models have included Lagrange multipliers as co-states. Often, the inclusion of these past multipliers as co-states is justified by the observation that past multipliers appear in the first-order conditions of the Ramsey problem. Our work provides a formal proof that, under standard assumptions, co-state past multipliers deliver the optimal solution in a general framework, encompassing a larger class of models with *forward-looking* constraints where, typically, optimal solutions are *constrained efficient solutions*.

We have commented at length on the relationship with the promised-utility approach in sub-sections 3.3. This approach has been widely used in macroeconomics[33]. Some applications are by Kocherlakota (1995) in a model with participation constraints, and Cronshaw and Luenberger (1994) in a dynamic game. Moreover, Kydland and Prescott (1980), Chang (1998) and Phelan and Stacchetti (2001) study Ramsey equilibria using promised *marginal* utility as a co-state variable, and they note the analogy of their approach with promised utility.

One advantage of the promised-utility approach is that it naturally allows for the characterization of all feasible paths (not only the constrained-efficient) and it naturally applies to models with private information or models with multiple solutions. However, the initial advantages of the promised-utility approach have mostly vanished. For example, Sleet and Yeltekin (2010) and Mele (2014) have extended our approach to address moral hazard problems and Ábrahám *et al.* (2017) study a risk-sharing partnership with intertemporal participation and moral hazard constraints.

Many applications of our approach can be found in the literature and we mention some in the Introduction. Nevertheless, it is beyond the scope of this paper to discuss in detail this expanding literature, which seems to be testimony to the convenience of using our approach, especially in models with natural state variables such as capital (as in Example 1) or debt (as in Example 2).

Perhaps it is most interesting that the approach here can be used as an intermediate step in solving models that go beyond the pure formulation of $\mathbf{PP}_\mu$. For example, a second generation of models considers *endogenous* participation constraints, as in the non-market exclusion models of Cooley et al. (2004), Marimon and Quadrini (2011), and Ferrero and Marcet (2005). In these models, the functions h that appear in the incentive constraints are endogenous; they depend on the optimal or equilibrium solution, and the approach of this paper is often used as an intermediate step, defining the underlying contracts. This allows the study of problems where the outside option is determined in equilibrium as in models of debt renegotiation and long-term contracts. Furthermore, the work of Debortoli and Nunes (2010) extends our approach to studying models of partial commitment and political economy.

[33] Ljungqvist and Sargent (2012) provide an excellent introduction and references to most of this recent work.

7 Concluding remarks

We have shown that a large class of problems with *forward-looking* constraints can be analysed using an equivalent saddle-point problem. This saddle-point problem obeys a saddle-point functional equation (**SPFE**) which is a version of the Bellman equation. The approach works for a very large class of models with incentive constraints: intertemporal enforcement constraints, intertemporal Euler equations in optimal policy and regulation design, etc. We provide a unified framework for the analysis of 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 *forward-looking* constraints, starting from pre-specified initial conditions. This simple representation also provides economic insight into the analysis of various contractual problems. For example, with intertemporal participation constraints it shows how the (Benthamite) social planner changes the weights assigned to different agents in order to keep them within the social contract; in Ramsey optimal problems it shows the cost of commitment to the benevolent government.

This paper provides the first complete account of the basic theory of *recursive contracts*. We have already presented most of the elements of the theory in our previous work (in particular, Marcet and Marimon (1988, 1999 & 2011), which has allowed others to build on it. Many applications are already found in the literature, showing the convenience of our approach, especially when natural state variables are present, or when our co-state variable μ plays a key rule in determining constrained efficiency *wedges* or pricing contracts. Similarly, extensions are already available, encompassing a wider set of problems than those considered here. Our sufficiency result is very general, although it requires checking that a *consistency condition* is satisfied, when there are *forward-looking* constraints. We also show that this condition is satisfied if there is a locally unique allocation that solves the **SPFE**; a wide range of applications to economics have this feature[34].

Authors' Affiliations:

Albert Marcet, Institut d'Analisi Economica CSIC, ICREA, UAB - Barcelona GSE, CEPR and MOVE

E-mail: a.marcet@iae.csic.es

and

Ramon Marimon, European University Institute, UPF - Barcelona GSE, NBER and CEPR,

Via delle Fontanelle 18, I-50014 San Domenico di Fiesole (FI) - Italy;

Telephone: +39-055-4685911,

E-mail: ramon.marimon@eui.eu

[34] Cole and Kubler (2012) provide a generalization to the non-uniqueness case for a restricted class of models. The recent work of Marimon and Werner (2017) follows the general framework presented here more closely and applying their extension of the Envelope Theorem generalizes our approach to the non-differentiable case (see footnote 23).

APPENDIX

Appendix A. The ∞-dimensional formulation and proofs of Theorem 2 and Proposition 1.

We first describe the *infinite-dimensional* formulation of the problems we study, which is used for Theorems 1 and 2. The underlying uncertainty takes the form of 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 $(s_0, ..., s_t) \in S_t$, $\mathcal{S}_t$ the σ-algebra of events of s^t and $\{s_t\}_{t=0}^{\infty} \in S_\infty$, with $\mathcal{S}$ the corresponding σ-algebra. 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 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 $\mathcal{A} = \{\boldsymbol{a} : \forall t \geq 0,\, a_t : S_t \to A \text{ and } a_t \in \mathcal{L}_\infty^m(S_t, \mathcal{S}_t, P),\, \}$, where $\mathcal{L}_\infty^m(S_t, \mathcal{S}_t, P)$ denotes the space of m-valued, essentially bounded, $\mathcal{S}_t$-measurable functions. The corresponding endogenous state variables are elements of $\mathcal{X} = \{\boldsymbol{x} : \forall t \geq 0,\, x_t \in \mathcal{L}_\infty^n(S_t, \mathcal{S}_t, P)\}$.

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

$$f_{(x,\mu.s)}(\boldsymbol{a}) = \mathrm{E}_0 \sum_{j=0}^{k} \sum_{t=0}^{N_j} \beta^t \mu^j h_0^j(x_t,\, a_t,\, s_t).$$

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

$$g(\boldsymbol{a})_t^j = \mathrm{E}_t \left[\sum_{n=1}^{N_j+1} \beta^n h_0^j(x_{t+n},\, a_{t+n},\, s_{t+n}) \right] + h_1^j(x_t,\, a_t,\, s_t).$$

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

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

Then the $\mathbf{PP}_\mu$ 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 as $\boldsymbol{x}^*$. When the solution exists we define the value function of $\mathbf{PP}_\mu$ as

$$V_\mu(x,s) = f_{(x,\mu.s)}(\boldsymbol{a}^*). \tag{38}$$

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

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

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

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 $\mathcal{B}'$, while these constraints form part of the objective function of $\mathbf{SPP}_{\mu}$.

Proof of Theorem 2: 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).

The point $(\boldsymbol{a}^*, \gamma^*)$ is a solution of $\mathbf{SPP}_{\mu}$ if and only if

$$f(\boldsymbol{a}^*) + \gamma g(\boldsymbol{a}^*)_0 \geq f(\boldsymbol{a}^*) + \gamma^* g(\boldsymbol{a}^*)_0 \geq f(\mathbf{a}) + \gamma^* g(\boldsymbol{a}^*)_0$$

The first inequality 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, since $\gamma = 0$ is a possible value, 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 exists $\tilde{\boldsymbol{a}} \in \mathcal{B}(x,s)$ 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 second inequality of the saddle-point condition for $\mathbf{SPP}_{\mu}$.

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

To prove the statement in the last sentence of the theorem, note that if $(\boldsymbol{a}^*, \gamma^*)$ is a solution to $\mathbf{SPP}_{\mu}$ in state (x,s), then

$$\begin{aligned}
SV(x,\mu,s) &= \mu h_0(x_0, a_0^*, s_0) + \gamma^* h_1(x_0, a_0^*, s_0) \\
&\qquad + \beta\ \mathrm{E}_0 \sum_{j=0}^{l} \varphi^j(\mu, \gamma^*) \sum_{t=0}^{N_j} \beta^t\ h_0^j(x_{t+1}^*, a_{t+1}^*, s_{t+1}) \\
&= \sum_{j=0}^{l} \mu^j \left[h_0^j(x_0, a_0^*, s_0) + \beta\ \mathrm{E}_0 \sum_{t=0}^{N_j} \beta^t\ h_0^j(x_{t+1}^*, a_{t+1}^*, s_{t+1}) \right] \\
&+ \sum_{j=0}^{l} \gamma^{*j} \left[h_1^j(x_0, a_0^*, s_0) + \beta\ \mathrm{E}_0 \sum_{t=0}^{N_j} \beta^t\ h_0^j(x_{t+1}^*, a_{t+1}^*, s_{t+1}) \right] \\
&= \mathrm{E}_0 \left[\sum_{j=0}^{l} \sum_{t=0}^{N_j} \beta^t \mu^j h_0^j(x_t^*, a_t^*, s_t) \mid s \right] \\
&= V_{\mu}(x,s),
\end{aligned}$$

where the first equality simply rearranges terms, the second follows from the standard Kuhn-Tucker condition and the last from the first statement in the Theorem ■

Proof of Proposition 1 Let $\widehat{S}_1 \subset S$ be the set such that if $s_1 \in \widehat{S}_1$ then $V_{\mu_1^*}(x_1^*, s_1) > \mathrm{E}_1 \left[\sum_{j=0}^{l} \sum_{t=0}^{N_j} \beta^t \mu_1^{*j} h_0^j(x_{t+1}^*,\, a_{t+1}^*,\, s_{t+1})\right.$

We will show that $\widehat{S}_1$ has probability zero.

The constraints in $\mathbf{PP}_{\mu_1^*}$ are a subset of the constraints in $\mathbf{SPP}_\mu$. Therefore the continuation for $\boldsymbol{a}^*$, namely $\{a_t^*\}_{t=1}^\infty$, is feasible for $\mathbf{PP}_{\mu_1^*}$ with initial conditions (x_1^*, s_1). If $s_1 \in \widehat{S}_1$ there must exist a plan $\{\widehat{a}_t\}_{t=0}^\infty$ achieving a higher value for $\mathbf{PP}_{\mu_1^*}$ with initial conditions (x_1^*, s_1) than the value achieved by $\{a_t^*\}_{t=1}^\infty$ so that

$$\mathrm{E}_1 \left[\sum_{j=0}^{l} \mu_1^{j*} \sum_{t=0}^{N_j} \beta^t \; h_0^j(x_{t+1}^*,\, a_{t+1}^*,\, s_{t+1}) \mid s_1\right] < \mathrm{E}_1 \left[\sum_{j=0}^{l} \mu_1^{j*} \sum_{t=0}^{N_j} \beta^t \; h_0^j(\widehat{x}_t,\, \widehat{a}_t,\, s_{t+1}) \mid s_1\right].$$

If $\mathrm{Prob}(\widehat{S}_1) > 0$ we would have the inequality in

$$\begin{aligned} & \mu h_0(x_0,\, a_0^*,\, s_0) + \gamma_0^* h_1(x_0,\, a_0^*,\, s_0) + \beta\; \mathrm{E}_0 \left[\sum_{j=0}^{l} \mu_1^{j*} \sum_{t=0}^{N_j} \beta^t \; h_0^j(x_{t+1}^*,\, a_{t+1}^*,\, s_{t+1}) \mid s_0\right] \\ < \;& \mu h_0(x_0,\, a_0^*,\, s_0) + \gamma_0^* h_1(x_0,\, a_0^*,\, s_0) \\ & +\beta\; \mathrm{E}_0 \left[\mathrm{E}_1 \left[\sum_{j=0}^{l} \mu_1^{j*} \sum_{t=0}^{N_j} \beta^t \; h_0^j(\widehat{x}_{t+1},\, \widehat{a}_{t+1},\, s_{t+1}) \mid s_1 \in \widehat{S}_1\right] \mid s_0\right] \\ & +\beta\; \mathrm{E}_0 \left[\mathrm{E}_1 \left[\sum_{j=0}^{l} \mu_1^{j*} \sum_{t=0}^{N_j} \beta^t \; h_0^j(x_{t+1}^*,\, a_{t+1}^*,\, s_{t+1}) \mid s_1 \in S \backslash \widehat{S}_1\right] \mid s_0\right]. \end{aligned} \tag{39}$$

Since the forward-looking constraints in $\mathbf{SPP}_\mu$ with initial conditions (x, s) are the same as the constraints in $\mathbf{PP}_{\mu_1^*}$ with initial conditions (x_1^*, s_1), the plan $\left\{a_0^*, \{\widehat{a}_t\}_{t=1}^\infty\right\}$ is feasible for $\mathbf{SPP}_\mu$ conditional on realizations $s_1 \in \widehat{S}_1$. Then, the above inequality contradicts the '*sup*' property of $(\boldsymbol{a}^*, \gamma^*)$ as a saddle-point of $\mathbf{SPP}_\mu$ with initial conditions (x, s). Therefore $\mathrm{Prob}(\widehat{S}_1) = 0$ or, equivalently, $V_{\mu_1^*}(x_1^*, s_1) \leq \mathrm{E}_1 \left[\sum_{j=0}^{l} \sum_{t=0}^{N_j} \beta^t \mu_1^{*j} h_0^j(x_{t+1}^*,\, a_{t+1}^*,\, s_{t+1}) \mid s_1\right]$ a.s.

Using, again, the fact that the continuation of a feasible sequence for $\mathbf{SPP}_\mu$ satisfies the constraints of $\mathbf{PP}_{\mu_1^*}$, we have $V_{\mu_1^*}(x_1^*, s_1) \geq \mathrm{E}_1 \left[\sum_{j=0}^{l} \sum_{t=0}^{N_j} \beta^t \mu_1^{*j} h_0^j(x_{t+1}^*,\, a_{t+1}^*,\, s_{t+1}) \mid s_1\right]$.

Therefore, $\{a_t^*\}_{t=1}^\infty$ solves $\mathbf{PP}_{\mu_1^*}$ with initial conditions (x_1^*, s_1) a.s.■

Appendix B. Supporting results for the proofs of Theorems 3 and 4.

Some properties of SV and SPFE

Lemma 1A. Assume $\mathbf{SPP}_\mu$ has a solution at (x, s) with value $SV(x, \mu, s)$, for $x \in X$ and $\mu \in R_+^{l+1}$. Then

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

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

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

Proof: *i)* To simplify notation, denote the solution of $\mathbf{SPP}_\mu$ at (x, s) by $(\boldsymbol{a}_\mu^*, \gamma_\mu^*)$ and note that, by the definition of f, given any $\boldsymbol{a}$, $\mu, \mu' \in R_+^{l+1}$ and scalars λ, λ' we have

$$f_{(x, \lambda\mu + \lambda'\mu', s)}(\boldsymbol{a}) = \lambda f_{(x,\mu,s)}(\boldsymbol{a}) + \lambda' f_{(x,\mu',s)}(\boldsymbol{a}) \tag{40}$$

and, in particular, that $f_{(x,\lambda\mu,s)}(\boldsymbol{a}) = \lambda f_{(x,\mu,s)}(\boldsymbol{a})$.

To prove convexity note that given any $\mu, \mu' \in R_+^{l+1}$ and a scalar $\lambda \in (0,1)$, we have

$$\begin{aligned}
SV(x, \lambda\mu + (1-\lambda)\mu', s) &= \lambda f_{(x,\mu,s)}(\boldsymbol{a}^*_{\lambda\mu+(1-\lambda)\mu'}) + (1-\lambda) f_{(x,\mu',s)}(\boldsymbol{a}^*_{\lambda\mu+(1-\lambda)\mu'}) \\
&+ \gamma^*_{\lambda\mu+(1-\lambda)\mu'} g_0(\boldsymbol{a}^*_{\lambda\mu+(1-\lambda)\mu'}) \\
&\leq \lambda[f_{(x,\mu,s)}(\boldsymbol{a}^*_{\lambda\mu+(1-\lambda)\mu'}) + \gamma^*_\mu g_0((\boldsymbol{a}^*_{\lambda\mu+(1-\lambda)\mu'})] \\
&+ (1-\lambda)[f_{(x,\mu',s)}(\boldsymbol{a}^*_{\lambda\mu+(1-\lambda)\mu'}) + \gamma^*_{\mu'} g_0(\boldsymbol{a}^*_{\lambda\mu+(1-\lambda)\mu'})] \\
&\leq \lambda[f_{(x,\mu,s)}(\boldsymbol{a}^*_\mu) + \gamma^*_\mu g_0((\boldsymbol{a}^*_\mu)] \\
&+ (1-\lambda)[f_{(x,\mu',s)}(\boldsymbol{a}^*_{\mu'}) + \gamma^*_{\mu'} g_0((\boldsymbol{a}^*_{\mu'})] \\
&= \lambda SV(x,\mu,s) + (1-\lambda) SV(x,\mu',s),
\end{aligned}$$

where the first equality follows from (40), the first inequality follows from from the fact that $\gamma^*_{\lambda\mu+(1-\lambda)\mu'}$ minimizes $\mathbf{SPP}_{\lambda\mu+(1-\lambda)\mu'}$ and the second from the fact that $\boldsymbol{a}^*_\mu$ and $\boldsymbol{a}^*_{\mu'}$ maximize $\mathbf{SPP}_\mu$ and $\mathbf{SPP}_{\mu'}$, respectively.

To prove homogeneity of degree one, fix a scalar $\lambda > 0$. Then, using (40) and the fact that $\boldsymbol{a}^*_{\lambda\mu}$ and $\boldsymbol{a}^*_\mu$ are maximal elements attaining $SV(x,\lambda\mu,s)$ and $SV(x,\mu,s)$ respectively:

$$\begin{aligned}
SV(x,\lambda\mu,s) &= f_{(x,\lambda\mu,s)}(\boldsymbol{a}^*_{\lambda\mu}) \geq f_{(x,\lambda\mu,s)}(\boldsymbol{a}^*_\mu) \\
= \lambda f_{(x,\mu,s)}(\boldsymbol{a}^*_\mu) &= \lambda SV(x,\mu,s) \geq \lambda f_{(x,\mu,s)}(\boldsymbol{a}^*_{\lambda\mu}) \\
= f_{(x,\lambda\mu,s)}(\boldsymbol{a}^*_{\lambda\mu}) &= SV(x,\lambda\mu,s).
\end{aligned}$$

The proofs of (*ii*) and (*iii*) are straightforward: in particular, *ii*) follows from applying the Theorem of the Maximum (Stokey, Lucas and Prescott,1989, Theorem 3.6) and *iii*) follows from the fact that the constraint sets are convex and the objective function concave ■

Lemma 2A: If the *saddle-point 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-hand side of **SPFE** at $(x,\ \mu,\ s)$, $(\widetilde{a}, \widetilde{\gamma})$ and $(\widehat{a}, \widehat{\gamma})$. Then by repeated application of the *saddle-point* min and max conditions:

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

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

Properties of convex homogeneous functions.

To simplify the exposition of these properties, let $F : R_+^m \to R$ be continuous and convex, satisfying $F(x) < \infty$ for some $x >> 0$. The *subdifferential 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 Lemma 1 and Proposition 2:

F1. *i)* $\partial F(y)$ is a closed and convex set; *ii)* if $y \in R^m_{++}$, $\partial F(y)$ is also non-empty and bounded, and *iii)* the correspondence $\partial F : R^m_+ \longrightarrow R^m$ is upper-hemi continuous.

F2. F 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.

F3. Lemma 3A (Euler's formula). If F is also homogeneous of degree one and $z \in \partial F(y)$, then $F(y) = yz$. Furthermore, for any $\lambda > 0, \partial F(\lambda y) = \partial F(y)$, i.e. the subdifferential is homogeneous of degree zero.

F4. Lemma 4A. (Kuhn-Tucker) x^* minimizes F on R^m_+ *if and only if* there is a $f(x^*) \in \partial F(x^*)$ such that: (i) $f(x^*) \geq 0$, and (ii) $x^* f(x^*) = 0$.

F5. If $F = \sum_{i=1}^m \alpha_i F^i$, where, for $i = 1, ..., m$, $\alpha_i > 0$ and $F^i : R^m_+ \rightarrow R$ is convex, then $\partial F(y) = \sum_{i=1}^m \alpha_i \partial F^i(y)$.

Facts **F1** and **F2** are well known and can be found in Rockafellar (1970): **F1(**i**)** follows immediately from the definition of the subdifferential (Ch. 23); **F1(**ii**)** from Theorem 23.4; **F1(**iii**)** from Theorem 24.4, and **F2** from Theorem 25.1. Similarly, **F5** follows from Theorem 23.8.

Proof of Lemma 3A: 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 $\lambda \in (0, 1)$, it results in $F(y) \leq yz$. Therefore $F(y) = yz$. To see that $\partial F(y)$ is homogeneous of degree zero note that, for any $\lambda > 0$,

$$\begin{aligned} \partial F(\lambda y) &= \left\{ z \in R^m \mid F(y') \geq F(\lambda y) + (y' - \lambda y)z \text{ for all } y' \in R^m_+ \right\} \\ &= \left\{ z \in R^m \mid F(\lambda y'') \geq F(\lambda y) + (\lambda y'' - \lambda y)z \text{ for all } y'' \in R^m_+ \right\} \\ &= \left\{ z \in R^m \mid F(y'') \geq F(y) + (y'' - y)z \text{ for all } y'' \in R^m_+ \right\} = \partial F(y) \ \blacksquare \end{aligned}$$

Proof of Lemma 4A: The proof is based on Rockafellar's (1981, Ch. 5) characterization of stationary points using subdifferential calculus (R81 in what follows). First, we prove *necessity*: let x^* minimize F on R^m_+. Since the constrained set is convex with a non-empty interior, x^* minimizes $F(x) - \lambda^* x$, where $\lambda^* \in R^m_+$ and $\lambda^{*j} = 0$ if $x^{*j} > 0$; otherwise x^* would not be a minimizer. By R81, Proposition 5A, $0 \in \partial\{F(x) - \lambda^* x\}$ and, since $\{x \in R^m_{++} | F(x) < \infty\} \neq \varnothing$, $\partial\{F(x) - \lambda^* x\} = \partial F(x) + \partial\{-\lambda^* x\}$ (R81, Theorem 5C); that is, there exists $f(x^*) \in \partial F(x^*)$ such that $f(x^*) - \lambda^* = 0$. Therefore, $f(x^*) \geq 0$ and $x^* f(x^*) - \lambda^* x^* = x^* f(x^*) = 0$.

To see *sufficiency*, note that since F is convex and $f(x^*) \in \partial F(x^*)$, for any $x \in R^m_+$, $F(x) - F(x^*) \geq (x - x^*) f(x^*)$, but given (i) and (ii) the inequality simplifies to $F(x) - F(x^*) \geq 0$ $\blacksquare$

Other supporting results

Proof of Lemma 1: Part $(\boldsymbol{i})$ follows from **F1 - F3**. In particular, **F3** implies that if $z \in \partial F(y)$ then $z \in \partial F(\lambda y)$. The saddle-point max inequality condition of part $(\boldsymbol{ii})$ (29) is the same as the max saddle-point condition of **SPFE** expressed with its Euler representation. Since by $(\boldsymbol{i})$ W always has at least one Euler representation, the proof of (29) is immediate. To see the min inequality of part (***ii***), begin by rewriting the first inequality of $\Psi_W(x,\,\mu,\,s)$, (25, as:

$$\gamma h_1(x,\,a^*,\,s)+\beta \mathrm{E}\left[W(\ell(x,\,a^*,\,s'),\,\varphi(\mu,\,\gamma),\,s')|\,s\right] \geq \gamma^* h_1(x,\,a^*,\,s)+\beta \mathrm{E}\left[W(\ell(x,\,a^*,\,s'),\,\varphi(\mu,\,\gamma^*),\,s')|\,s\right].$$

Then, let

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

By **F5,**

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

and it follows from **F4** (Lemma 4A) that the Kuhn-Tucker conditions (30) and (31) are necessary and sufficient. ■

Appendix C. Proofs of Theorems 3 and 4.

Proof of Theorem 3: We start by proving that if (a_0^*, γ^*) is the period zero solution of $\mathbf{SPP}_\mu$ at $(x,\,s)$ then it is a saddle-point of **SPFE** at (x, μ, s).

First, we show that, given γ^*, a_0^* solves (26) when $W = SV$. Take any $\widetilde{a} \in A$ such that $p(x, \widetilde{a}, s) \geq 0$. Consider the sequence obtained by starting at $\widetilde{a}$ and then continuing to the optimal solution of $\mathbf{PP}_{\mu_1^*}$ from $t = 1$ onwards. To properly express this we introduce some notation. Let the shift operator $\sigma : S^{t+1} \to S^t$ be given by $\sigma(s^t) \equiv \sigma(s_0, s_1, \ldots, s_t) = (s_1, s_2..., s_t)$, and – denoting $(\boldsymbol{a}^*(x, \mu, s), \gamma^*(x, \mu, s))$ a solution to $\mathbf{SPP}_\mu$ at $(x,\,s)$ – let the solution plan following a deviation $\widetilde{a}$ have the following representation:

$$\begin{aligned}\widetilde{a}_0(x, \mu, s) &= \widetilde{a} \text{ and}\\ \widetilde{a}_t(x, \mu, s)(s^t) &= a_{t-1}^*(\widetilde{x}_1, \mu_1^*(x, \mu, s),\ s_1)(\sigma(s^t)) \text{ for all } t > 0,\end{aligned}$$

where $\widetilde{x}_1 = \ell(x,\,\widetilde{a},\,s_1)$. Theorem 2 and the definition of $\mathbf{PP}_{\mu_1^*}$ imply that

$$\mathrm{E}\left[SV(\widetilde{x}_1,\,\mu_1^*,\,s_1)|s\right] = \mathrm{E}\left[V_{\mu_1^*}(\widetilde{x}_1,\,s_1)|s\right] = \mathrm{E}_0 \sum_{j=0}^{l} \mu_1^{*j} \sum_{t=0}^{N_j} \beta^t\ h_0^j(\widetilde{x}_{t+1},\,\widetilde{a}_{t+1},\,s_{t+1}). \tag{41}$$

This equality also works for $\boldsymbol{a}^*$ instead of $\widetilde{\boldsymbol{a}}$. Using this and the fact that the sequence $\widetilde{\boldsymbol{a}}$ is feasible for $\mathbf{SPP}_\mu$ (as the proof of Proposition 1) and that $\boldsymbol{a}^*$ solves the *sup* part of $\mathbf{SPP}_\mu$ we have

$$\begin{aligned}&\mu h_0(x,\,\widetilde{a},\,s) + \gamma^* h_1(x,\,\widetilde{a},\,s) + \beta \mathrm{E}\left[SV(\widetilde{x}_1,\,\mu_1^*,\,s_1)|\,s\right]\\ \leq\ &\mu h_0(x,\,a_0^*,\,s) + \gamma^* h_1(x,\,a_0^*,\,s) + \beta \mathrm{E}_0 \sum_{j=0}^{l} \mu_1^{*j} \sum_{t=0}^{N_j} \beta^t\ h_0^j(x_{t+1}^*,\,a_{t+1}^*,\,s_{t+1})\\ =\ &\mu h_0(x,\,a_0^*,\,s) + \gamma^* h_1(x,\,a_0^*,\,s) + \beta \mathrm{E}\left[SV(x_1^*,\,\mu_1^*,\,s_1)|\,s\right].\end{aligned}$$

This proves that a_0^* solves (26) when $W = SV$.

A similar argument shows that γ^* solves (25). Given any $\widetilde{\gamma} \in R_+^{l+1}$, Theorem 2 implies that

$$\begin{aligned}\beta \mathrm{E}\left[SV(x_1^*, \varphi(\mu, \widetilde{\gamma}), s_1)| s\right] &= \beta \mathrm{E}\left[V_{\varphi(\mu, \widetilde{\gamma})}(x_1^*, s_1)|s\right] \\ &\geq \mathrm{E}\left[\sum_{j=0}^{l} \varphi(\mu, \widetilde{\gamma})^j \sum_{t=0}^{N_j} \beta^t \; h_0^j(x_{t+1}^*, a_{t+1}^*, s_{t+1}) \mid s\right],\end{aligned}$$

where the inequality follows from the fact that the continuation of $\boldsymbol{a}^*$ is feasible but not necessarily optimal for $\mathbf{PP}_{\varphi(\mu, \widetilde{\gamma})}$. Using this and the fact that γ^* solves the min part of $\mathbf{SPP}_\mu$, we have

$$\begin{aligned}&\mu h_0(x, a_0^*, s) + \widetilde{\gamma} h_1(x, a_0^*, s) + \beta \mathrm{E}\left[SV(x_1^*, \varphi(\mu, \widetilde{\gamma}), s')| s\right] \\ \geq \; &\mu h_0(x, a_0^*, s) + \gamma^* h_1(x, a_0^*, s) + \beta \mathrm{E}\left[SV(x_1^*, \varphi(\mu, \gamma^*), s_1)| s\right].\end{aligned}$$

This proves that $(a_0^*(x, \mu, s), \gamma^*(x, \mu, s)) \in \Psi_{SV}(x, \mu, s)$. Finally, using the definition of SV in (20) we have

$$SV(x, \mu, s) = \mu h_0(x, a_0^*, s) + \gamma^* h_1(x, a_0^*, s) + \beta \; \mathrm{E}\left[SV(x_1^*, \varphi(\mu, \gamma^*), s')| s\right]. \tag{42}$$

Therefore SV satisfies **SPFE** ■

Proof of Theorem 4: Let $(\mathbf{a}^*, \boldsymbol{\gamma}^*)_{(x,\mu,s)}$ be generated by the *saddle-point policy correspondence* Ψ_W that is, $(a_t^*, \gamma_t^*) \in \Psi_W(x_t^*, \mu_t^*, s_t)$ for every (t, s_t). We first show that a^* satisfies the feasibility conditions of $\mathbf{SPP}_\mu$ at (x, s). The technological constraints (2) are satisfied by the definition of Ψ_W, so we only need to show that the forward-looking constraints (3) are satisfied for $t \geq 1$. By the slackness condition (30) of Lemma 1 *ii)* it follows that

$$h_1^j(x_t^*, a_t^*, s_t) + \beta \; \mathrm{E}\left[\omega^j(x_{t+1}^*, \mu_{t+1}^*, s_{t+1})| s_t\right] \geq 0. \tag{43}$$

By definition, for $j = k+1, ..., l, \varphi^j(\mu_t^*, \gamma_t^*) = \gamma_t^{*j}$, the j^{th} component of $W(x_t^*, \mu_t^*, s_t)$ is $\mu_t^j h_0^j(x_t^*, a^*, s_t) + \gamma_t^{*j}\left[h_1^j(x_t^*, a^*, s_t) + \beta \; \mathrm{E}\left[\omega^j(x_{t+1}^*, \mu_{t+1}^*, s_{t+1})| s_t\right]\right] = \mu_t^j h_0^j(x_t^*, a^*, s_t)$. This means that the partial derivative of W with respect to μ^j exists and it is given by $\partial_{\mu^j} W(x_t^*, \mu_t^*, s_t) = \omega^j(x_t^*, \mu_t^*, s_t) = h_0^j(x_t^*, a^*, s_t)$. Therefore (43) becomes the feasibility condition

$$h_1^j(x_t^*, a_t^*, s_t) + \beta \; \mathrm{E}\left[h_0^j(x_{t+1}^*, a_{t+1}^*, s_{t+1})| s_t\right] \geq 0.$$

For $j = 0, ..., k$, we iterate on (27) and, given the transversality condition (28), we obtain: $\omega_t^j(x_t^*, \mu_t^*, s_t) = \mathrm{E}\left[\sum_{n=0}^{\infty} \beta^t h_0^j(x_{t+n}^*, a_{t+n}^*, s_{t+n})| s_t\right]$ and (43) results in:

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

The two inequalities above correspond to the *forward looking constraints* (3) when they take the form $N_j = 0$, for $j = k+1, \ldots, l$ and $N_j = \infty$, for $j = 0, \ldots, k$, respectively. Now, to see that solutions to the **SPFE** are, in fact, solutions to $\mathbf{SPP}_\mu$ we need to show that the following saddle-point condition is satisfied

$$\begin{aligned}
&\mu h_0(x_0,\, a_0^*,\, s_0) + \gamma h_1(x_0,\, a_0^*,\, s_0) + \beta \ \mathrm{E}_0 \sum_{j=0}^{l} \varphi^j(\mu,\, \gamma) \sum_{t=0}^{N_j} \beta^t \ h_0^j(x_{t+1}^*,\, a_{t+1}^*,\, s_{t+1}) \\
\geq \quad &\mu h_0(x_0,\, a_0^*,\, s_0) + \gamma_0^* h_1(x_0,\, a_0^*,\, s_0) + \beta \ \mathrm{E}_0 \sum_{j=0}^{l} \varphi^j(\mu,\, \gamma_0^*) \sum_{t=0}^{N_j} \beta^t \ h_0^j(x_{t+1}^*,\, a_{t+1}^*,\, s_{t+1}) \\
\geq \quad &\mu h_0(x_0,\, a_0,\, s_0) + \gamma_0^* h_1(x_0,\, a_0,\, s_0) \\
&+\beta \ \mathrm{E}_0 \sum_{j=0}^{l} \varphi^j(\mu,\, \gamma_0^*) \sum_{t=0}^{N_j} \beta^t \ h_0^j(\ell(x_{t+1},\, a_{t+1},\, s_{t+1}),\, a_{t+1},\, s_{t+1}),
\end{aligned}$$

for any $\gamma \in R_+^{l+1}$ and $\{a_t\}_{t=0}^\infty$ satisfying the $\mathbf{SPP}_\mu$ constraints.

Let a program $\{\widetilde{a}_t\}_{t=0}^\infty$, and $\{\widetilde{x}_t\}_{t=0}^\infty$, given by $\widetilde{x}_0 = x, \widetilde{x}_{t+1} = \ell(\widetilde{x}_t,\, \widetilde{a}_t,\, s_{t+1})$, satisfy the constraints of $\mathbf{SPP}_\mu$ with initial condition (x, s). Since

$$\sum_{t=0}^{N_j} \beta^t \ h_0^j(x_{t+1}^*,\, a_{t+1}^*,\, s_{t+1}) = \omega(\ell(x,\, a_0^*,\, s_1),\, \varphi(\,\mu,\, \gamma_0^*),\, s_1),$$

the min part of $\mathbf{SPP}_\mu$ (i.e. the first inequality above) is given by the Kuhn-Tucker conditions (30) and (31). To prove the max part requires more work and we first introduce some additional notation. Given $x \in R^{l+1}$, let $I^k x^j = x^j$ if $j = 0, ..., k$ and $I^k x^j = 0$ if $j = k+1, ..., l$ (i.e. the projection in R^{k+1}). Let $\mu_1^* = \varphi\left(\mu, \gamma_0^*\right), \widetilde{\mu}_2^* = \varphi(\mu_1^*, \gamma_1^*(\widetilde{x}_1))$[35] and, for $t > 1$ $\widetilde{\mu}_{t+1}^* = \varphi\left(\widetilde{\mu}_t^*, \gamma_t^*(\widetilde{x}_t)\right)$; that is, $\widetilde{\mu}_t^*$ is the co-state for the deviation plan. In what follows, we proceed by iteration of the **SPFE** (max) inequality, (29), and expansion of the value function, according to (33); in particular the inequalities (44), (46) and (49) apply the inequality (29), and the equalities (45) and (47) apply the equality (33), while equality (48) simply rearranges terms and (50) uses the transversality condition, $\lim_{T \longrightarrow \infty} \beta^T W = 0$, to conclude the proof of the max part of **SPP**, with the left-hand side of (44) being greater or equal to (51):

[35] We also simplify notation by writing simply $\gamma_1^*(\widetilde{x}_1)$ instead of $\gamma_1^*(\widetilde{x}_1, \mu_1^*, s_1)$.

$$
\begin{aligned}
& \mu h_0(x,\, a_0^*,\, s) + \gamma_0^* h_1(x,\, a_0^*,\, s) + \beta\varphi\left(\mu,\gamma_0^*\right) \mathrm{E}\left[\omega(\ell(x,\, a_0^*,\, s_1),\, \varphi\left(\mu,\gamma_0^*\right),\, s_1)|\, s\right] \\
\geq\;& \mu h_0(x,\, \widetilde{a}_0,\, s) + \gamma_0^* h_1(x,\, \widetilde{a}_0,\, s) + \beta\varphi\left(\mu,\gamma_0^*\right) \mathrm{E}\left[\omega(\ell(x,\, \widetilde{a}_0,\, s_1),\, \varphi\left(\mu,\gamma_0^*\right),\, s_1)|\, s\right] && (44) \\
=\;& \mu h_0(x,\, \widetilde{a}_0,\, s) + \gamma_0^* h_1(x,\, \widetilde{a}_0,\, s) \\
& +\beta \mathrm{E}\mu_1^* \left[h_0(\widetilde{x}_1,\, a_1^*(\widetilde{x}_1),\, s_1) + \beta \mathrm{E}\left[I^k \omega(\ell(\widetilde{x}_1,\, a_1^*(\widetilde{x}_1),\, s_2), \varphi\left(\mu_1^*, \gamma_1^*(\widetilde{x}_1)\right),\, s_2)|\, s_1\right] |\, s\right] \\
& +\beta \mathrm{E}\gamma_1^*(\widetilde{x}_1)\left[h_1(\widetilde{x}_1,\, a_1^*(\widetilde{x}_1),\, s_1) + \beta \mathrm{E}\left[\omega(\ell(\widetilde{x}_1,\, a_1^*(\widetilde{x}_1),\, s_2), \varphi\left(\mu_1^*, \gamma_1^*(\widetilde{x}_1)\right),\, s_2)|\, s_1\right] |\, s\right] && (45) \\
\geq\;& \mu h_0(x,\, \widetilde{a}_0,\, s) + \gamma_0^* h_1(x,\, \widetilde{a}_0,\, s) \\
& +\beta \mathrm{E}\mu_1^* h_0(\widetilde{x}_1,\, \widetilde{a}_1,\, s_1) + \beta \mathrm{E}\left[I^k \omega(\ell(\widetilde{x}_1, \widetilde{a}_1,\, s_2), \varphi\left(\mu_1^*, \gamma_1^*(\widetilde{x}_1)\right),\, s_2)|\, s_1\right] |\, s] \\
& +\beta \mathrm{E}\gamma_1^*(\widetilde{x}_1)\left[h_1(\widetilde{x}_1,\, \widetilde{a}_1,\, s_1) + \beta \mathrm{E}\left[\omega(\ell(\widetilde{x}_1, \widetilde{a}_1,\, s_2), \varphi\left(\mu_1^*, \gamma_1^*(\widetilde{x}_1)\right),\, s_2)|\, s_1\right] |\, s\right] && (46) \\
=\;& \mu\left[h_0(x,\, \widetilde{a}_0,\, s) + \beta \mathrm{E}\left[I^k h_0(\widetilde{x}_1,\, \widetilde{a}_1,\, s_1)|\, s\right]\right] \\
& +\gamma_0^* \left[h_1(x,\, \widetilde{a}_0,\, s) + \beta \mathrm{E}\left[h_0(\widetilde{x}_1,\, \widetilde{a}_1,\, s_1)|\, s\right]\right] + \beta I^k \mathrm{E}\left[\gamma_1^*(\widetilde{x}_1) h_1(\widetilde{x}_1,\, \widetilde{a}_1,\, s_1)|s\right] \\
& +\beta^2 \mathrm{E}\left[\varphi\left(\mu_1^*, \gamma_1^*(\widetilde{x}_1)\right) \omega(\ell(\widetilde{x}_1, \widetilde{a}_1,\, s_2), \varphi\left(\mu_1^*, \gamma_1^*(\widetilde{x}_1)\right),\, s_2)\right] && (47) \\
=\;& \mu\left[h_0(x,\, \widetilde{a}_0,\, s) + \beta I^k \mathrm{E}\left[h_0(\widetilde{x}_1,\, \widetilde{a}_1,\, s_1) + \beta \mathrm{E}\left[h_0(\widetilde{x}_2,\, a_2^*(\widetilde{x}_2),\, s_2)|s_1\right] |\, s\right]\right] \\
& +\gamma_0^* h_1(x,\, \widetilde{a}_0,\, s) + \beta \mathrm{E}\left[h_0(\widetilde{x}_1,\, \widetilde{a}_1,\, s_1) + \beta I^k \mathrm{E}\left[h_0(\widetilde{x}_2,\, a_2^*(\widetilde{x}_2),\, s_2)|s_1\right] |\, s\right] \\
& +\beta \mathrm{E}\gamma_1^*(\widetilde{x}_1)\left[h_1(\widetilde{x}_1,\, \widetilde{a}_1,\, s_1) + \beta \mathrm{E}\left[h_0(\widetilde{x}_2,\, a_2^*(\widetilde{x}_2),\, s_2)|s_1\right] |s\right] + \beta^2 I^k \mathrm{E}\left[\gamma_2^*(\widetilde{x}_2) h_1(\widetilde{x}_2,\, a_2^*(\widetilde{x}_2),\, s_2)|s\right] \\
& +\beta^3 \mathrm{E}\varphi\left(\mu_2^*, \gamma_2^*(\widetilde{x}_2)\right) \left[\omega(\ell(\widetilde{x}_2,\, a_2^*(\widetilde{x}_2),\, s_3),\, \varphi\left(\mu_2^*, \gamma_2^*(\widetilde{x}_2)\right),\, s_3)|\, s\right] && (48) \\
& \ldots \\
\geq\;& A_T \equiv \mu\left[h_0(x,\, \widetilde{a}_0,\, s) + \beta I^k \mathrm{E}\left[\sum_{t=0}^{T-1} \beta^t h_0(\widetilde{x}_{t+1},\, \widetilde{a}_{t+1},\, s_{t+1})|s\right]\right] \\
& +\gamma_0^* \left[h_1(x,\, \widetilde{a}_0,\, s) + \beta \mathrm{E}\left[h_0(\widetilde{x}_1,\, \widetilde{a}_1,\, s_1) + I^k \sum_{t=1}^{T-1} \beta^t h_0(\widetilde{x}_{t+1},\, \widetilde{a}_{t+1},\, s_{t+1})|s\right]\right] \\
& +\beta \mathrm{E}\left[\gamma_1^*(\widetilde{x}_1)\left[h_1(\widetilde{x}_1,\, \widetilde{a}_1,\, s_1) + \beta\left[h_0(\widetilde{x}_2,\, \widetilde{a}_2,\, s_2) + I^k \sum_{t=2}^{T-1} \beta^{t-1} h_0(\widetilde{x}_{t+1},\, \widetilde{a}_{t+1},\, s_{t+1})\right]\right] |s\right] \\
& \ldots \\
& +\beta^T \mathrm{E}\left[\gamma_T^*(\widetilde{x}_T) h_1(\widetilde{x}_T,\, \widetilde{a}_T,\, s_T)|s\right] \\
& +\beta^{T+1} \mathrm{E}\left[\varphi\left(\mu_T^*, \gamma_T^*(\widetilde{x}_T)\right) \omega(\ell(\widetilde{x}_T, \widetilde{a}_T,\, s_{T+1}), \varphi\left(\mu_T^*, \gamma_T^*(\widetilde{x}_T)\right),\, s_{T+1})|\, s\right], && (49) \\
\lim_{T \longrightarrow \infty} A_T \;=\;& \mu\left[h_0(x,\, \widetilde{a}_0,\, s) + \beta I^k \mathrm{E}\left[\sum_{t=0}^{\infty} \beta^t h_0(\widetilde{x}_{t+1},\, \widetilde{a}_{t+1},\, s_{t+1})|s\right]\right] \\
& +\gamma_0^* \left[h_1(x,\, \widetilde{a}_0,\, s) + \beta \mathrm{E}\left[h_0(\widetilde{x}_1,\, \widetilde{a}_1,\, s_1) + I^k \sum_{t=1}^{T-1} \beta^t h_0(\widetilde{x}_{t+1},\, \widetilde{a}_{t+1},\, s_{t+1})|s\right]\right] && (50) \\
=\;& \mu h_0(x,\, \widetilde{a}_0,\, s) + \gamma_0^* h_1(x,\, \widetilde{a}_0,\, s) \\
& +\beta\ \mathrm{E}\left[\sum_{j=0}^{l} \varphi^j(\mu,\, \gamma_0^*) \sum_{t=0}^{N_j} \beta^t\ h_0^j(\widetilde{x}_{t+1},\, \widetilde{a}_{t+1},\, s_{t+1})|s\right] \blacksquare && (51)
\end{aligned}
$$

Proof of Corollary to Theorem 4: *i*) Let $\omega^*(x,\,\mu\,,s) = \omega_0(x,\,\mu\,,s) \in \partial_\mu W(x,\,\mu\,,s)$ and, for all $t \geq 1$, let $\omega_t(x_t^*,\, \mu_t^*, s_t) = \omega^*(x_t^*,\, \mu_t^*, s_t) = \omega_{t-1}(x_t^*,\, \mu_t^*, s_t) \in \partial_\mu W(x_t^*,\, \mu_t^*,\, s_t)$, where $\omega_{t-1}(x_t^*,\, \mu_t^*, s_t)$, together with $(\mathbf{a}^*, \boldsymbol{\gamma}^*)_{(x,\mu,s)}$, satisfy (29) and (32) and, with $\omega_t(x_t^*,\, \mu_t^*, s_t)$, satisfy (34); that is, by recursive construction $\omega^*(x_t^*,\, \mu_t^*, s_t)$, together with $(\mathbf{a}^*)$, satisfies (27). *ii*) If W is differentiable in μ at every $(x_t^*,\, a_t^*,\, s_t)$ then $\partial_\mu W(x_t^*,\, \mu_t^*,\, s_t)$ is a singleton and for every $(x_t^*,\, a_t^*,\, s_t)$, $\omega_t(x_t^*,\, \mu_t^*,\, s_t)$ is uniquely defined; By Lemma 1 (*i*) and taking partial derivatives on the left- and right-hand sides of (34), one obtains the *intertemporal consistency condition* (27) as an *Euler equation* of the min component of the saddle-point. *iii*) By *i*) there are $\omega_t(x_t^*,\, \mu_t^*, s_t)$ satisfying (27) and iterating on it and using the transversality condition (28) we obtain $\omega_t^{*j}(x_t^*,\, \mu_t^*,\, s_t) = \mathrm{E}\left[\sum_{n=0}^{\infty} \beta^t h_0^j(x_{t+n}^*,\, a_{t+n}^*,\, s_{t+n}) |\, s_t\right]$, since $(\mathbf{a}^*)_{(x,\mu,s)}$ is uniquely determined, any $\omega_t(x_t^*,\, \mu_t^*, s_t)$, with $\{\omega_{t+n}(x_{t+n+1}^*,\, \mu_{t+n+1}^*, s_{t+n+1})\}_{n=0}^{\infty}$ supporting $\{a_{t+n}^*, \gamma_{t+n}^*\}$, satisfying (28), must also satisfy $\omega_t(x_t^*,\, \mu_t^*, s_t) = \omega^*(x_t^*,\, \mu_t^*, s_t)$ and, therefore, (27). Given this, to show that $(\mathbf{a}^*)_{(x,\mu,s)}$ is a solution to $\mathbf{SPP}_\mu$ at (x, s) we refer to the proof of Theorem 4 ■

Appendix D. Proof of Proposition 2.

We first state a lemma, of which we omit the proof since it is a simple application of the *Theorem of the Maximum* (e.g. Stokey *et el.* 1989, Theorem 3.6) to the min and max parts of the *saddle-point* inequalities.

Lemma 5A. Assume **A1-A4.** $\Psi_W : X \times R_+^{l+1} \times S \to A \times R_+^{l+1}$ is a compact-valued and upper-hemi-continuous correspondence.

Proof of Proposition 2(i): The proof is structured in three three steps. First, we show that **DPP** (36) has a solution. To see this, let $A(x, s) \equiv \{a \in A : \, p(x, a,\, s) \geq 0\}$ and

$$B_{(x,\mu,s)}(\gamma) \equiv \left\{a \in \; A(x,s) : \, x' = \ell(x,\, a,\, s'),\; p(x, a,\, s) \geq 0 \text{ and } (a, x') \text{ satisfy } (37)\right\}.$$

Given our assumptions, provided that $W \in \mathcal{M}_b$, $B_{(x,\mu,s)}(\gamma)$ is a closed and bounded convex set with a non-empty interior, and $\mu h_0(x,\, a,\, s) + \beta\; \mathrm{E}\left[\sum_{j=0}^{k} \mu^j \omega^j(x',\, \mu',\, s') |\, s\right]$ is continuous in a; therefore, the set $D_{(x,\mu,s)}(\gamma) = \arg\sup_{a \in B_{(x,\mu,s)}(\gamma)} \left\{\mu h_0(x,\, a,\, s) + \beta\; \mathrm{E}\left[\sum_{j=0}^{k} \mu^j \omega^j(x',\, \mu',\, s') |\, s\right]\right\}$ is non-empty. Furthermore, if $W \in \mathcal{M}_{bc}$ – and **IC** is satisfied – there is a *saddle-point* $(a^*, \gamma^*)(x, \mu, s; \gamma)$ which is characterized by Kuhn-Tucker necessary and sufficient conditions (see, for example, Rockafellar (1970) Theorem 28.2):

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

However, up to now we have taken μ' as given, but we need to show that it satisfies $\mu'^* = \mu + \gamma^*$, where γ^* is the Kuhn-Tucker multiplier of **DPP** at (x, μ, s). This requires a fixed-point argument. To this end, we show that the Kuhn-Tucker multipliers are uniformly bounded, which will give us compactness of the set of Kuhn-Tucker multipliers. This is second step: given $(x, \mu, s; \gamma)$, and $\mu' = \varphi(\mu, \gamma)$, let $\widetilde{a} \in A$ be the interior choice satisfying **SIC** for (x, μ, s) and γ^*. Then, using the same notation as in the proof of Theorem 4, the slackness condition $\gamma^* \left[h_1(x,\, a^*,\, s) + \beta\; \mathrm{E}\left[\omega(\ell(x,\, a^*,\, s'),\, \mu',\, s') |\, s\right]\right] = 0$

and **SIC,**

$$\begin{aligned} & \mu\Big[h_0(x,\,a^*,\,s)+\beta\;\mathrm{E}\Big[I^k\omega(\ell(x,\,a^*,\,s'),\,\mu',\,s')|\,s\Big] \\ & -\Big(h_0(x,\,\widetilde{a},\,s)+\beta\;\mathrm{E}\Big[I^k\omega(\ell(x,\,\widetilde{a},\,s'),\,\mu',\,s')|\,s\Big]\Big)\Big] \\ \geq\quad & \gamma^*\left[h_1(x,\,\widetilde{a},\,s)+\beta\;\mathrm{E}\left[\omega(\ell(x,\,\widetilde{a},\,s'),\,\mu',\,s')|\,s\right]\right]\geq\varepsilon\left\|\gamma^*\right\|. \end{aligned}$$

If there is no uniform bound, then for any $\delta > 0$ there is a Kuhn-Tucker multiplier γ^* such that $\delta\left\|\gamma^*\right\| \geq \left\|\mu\right\|$, but in this case it must be that:

$$\begin{aligned} & \delta\frac{\mu}{\|\mu\|}\Big[h_0(x,\,a^*,\,s)+\beta\;\mathrm{E}\Big[I^k\omega(\ell(x,\,a^*,\,s'),\,\mu',\,s')|\,s\Big] \\ & -\Big(h_0(x,\,\widetilde{a},\,s)+\beta\;\mathrm{E}\Big[I^k\omega(\ell(x,\,\widetilde{a},\,s'),\,\mu',\,s')|\,s\Big]\Big)\Big] \\ \geq\quad & \frac{\mu}{\|\gamma^*\|}\Big[h_0(x,\,a^*,\,s)+\beta\;\mathrm{E}\Big[\omega^j(\ell(x,\,a^*,\,s'),\,\mu',\,s')|\,s\Big] \\ & -\Big(h_0(x,\,\widetilde{a},\,s)+\beta\;\mathrm{E}\Big[I^k\omega(\ell(x,\,\widetilde{a},\,s'),\,\mu',\,s')|\,s\Big]\Big)\Big] \\ \geq\quad & \frac{\gamma^*}{\|\gamma^*\|}\left[h_1(x,\,\widetilde{a},\,s)+\beta\;\mathrm{E}\left[\omega(\ell(x,\,\widetilde{a},\,s'),\,\mu',\,s')|\,s\right]\right]\geq\varepsilon, \end{aligned}$$

which is not possible for δ small enough, since all the terms in the main brackets are bounded. Therefore, there exists a $C > 0$ such that $\left\|\gamma^*\right\| \leq C\left\|\mu\right\|$. Third, we show that there is (a^*, γ^*) which is a solution to SPFE$_{(x,\mu,s)}$. Let $G_\mu = \left\{\gamma \in R_+^{l+1} : \|\gamma\| \leq C\left\|\mu\right\|\right\}$. As we have just seen, there is no loss of generality in restricting the choice of Kuhn-Tucker multipliers to G_μ if we can show that a solution exists with this restriction. For any $\gamma \in \mathcal{R}_+^{l+1}$, let

$$\begin{aligned} H_{(x,\mu,s)}(a,\gamma;\widehat{\gamma}) \quad\equiv\quad & h_0(x,\,a,\,s)+\beta\;E\left[I^k\omega(\ell(x,\,a,\,s'),\,\varphi(\mu,\widehat{\gamma}),\,s')|\,s\right] \\ & +\gamma\left[h_1(x,\,a,\,s)+\beta\;\mathrm{E}\left[\omega(\ell(x,\,a,\,s'),\,\varphi(\mu,\widehat{\gamma}),\,s')|\,s\right]\right], \end{aligned}$$

which allows us to define the set of *saddle-points* $(a^*, \gamma^*)(x, \mu, s; \widehat{\gamma})$ as

$$\begin{aligned} SP_{(x,\mu,s)}(\widehat{\gamma}) \quad=\quad & \{(a^*,\gamma^*)\in A(x,s)\times G_\mu:\forall(a,\gamma)\in A(x,s)\times G_\mu, \\ & H_{(x,\mu,s)}(a^*,\gamma;\widehat{\gamma})\geq H_{(x,\mu,s)}(a^*,\gamma^*;\widehat{\gamma})\geq H_{(x,\mu,s)}(a,\gamma^*;\widehat{\gamma})\Big\}. \end{aligned}$$

Since $\partial_\mu W(x, \cdot, s)$ is an upper-hemicontinuous correspondence (see **F1(*iii*)**) if $\widehat{\gamma}_n \longrightarrow \widehat{\gamma}$, then there is a subsequence $\omega(\ell(x,\,a,\,s'),\,\varphi(\mu,\widehat{\gamma}_n),\,s') \longrightarrow \omega(\ell(x,\,a,\,s'),\,\varphi(\mu,\widehat{\gamma}),\,s')$. Therefore, given our continuity assumptions, if $(a_n^*, \gamma_n^*; \widehat{\gamma}_n) \longrightarrow (a^*, \gamma^*; \widehat{\gamma})$ and, for all n, $(a_n^*, \gamma_n^*) \in SP_{(x,\mu,s)}(\widehat{\gamma}_n)$, then, for any $(a, \gamma) \in A(x, s) \times G_\mu$,

$$\begin{gathered} H_{(x,\mu,s)}(a_n^*,\gamma;\widehat{\gamma}_n)\geq H_{(x,\mu,s)}(a_n^*,\gamma_n^*;\widehat{\gamma}_n)\geq H_{(x,\mu,s)}(a,\gamma_n^*;\widehat{\gamma}_n) \\ \longrightarrow \\ H_{(x,\mu,s)}(a^*,\gamma;\widehat{\gamma})\geq H_{(x,\mu,s)}(a^*,\gamma^*;\widehat{\gamma})\geq H_{(x,\mu,s)}(a,\gamma^*;\widehat{\gamma}). \end{gathered}$$

In particular, $SP^2_{(x,\mu,s)}(\widehat{\gamma}) : G_\mu \longrightarrow G_\mu$ (i.e. the γ component of SP) is an upper-hemicontinuous, non-empty and compact-and-convex-valued correspondence, mapping a convex and compact set onto itself. Therefore, by *Kakutani's Fixed Point Theorem* (e.g. Mas-Colell et al. (1995), p.953) there is a fixed point $SP^2_{(x,\mu,s)}(\gamma^*) = \gamma^*$ and a corresponding *saddle-points* $(a^*, \gamma^*)(x, \mu, s; \gamma^*) \in SP_{(x,\mu,s)}(\gamma^*)$. If, in addition, strict concavity **A6s** is assumed, then $SP^1_{(x,\mu,s)}(\gamma^*)$ is a singleton a^*■

Proof of Proposition 2(ii): By assumption – or if the assumptions are satisfied, by Proposition 2 (i) – there is a solution to the *saddle-point* Bellman equation. Therefore, we only need to show that $T^*W \in \mathcal{M}$. By assumptions **A2**, **A3** and **A5,** and the definition of φ, the correspondences $\Gamma : X \to X$ and $\Phi : \mathcal{R}_+^{l+1} \to \mathcal{R}_+^{l+1}$ 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 the *Theorem of the Maximum* and assumption **A1b** (see Stokey, et al. (1989), Lemma 9.5) it follows that $T^*W(\cdot, \cdot, s)$ is continuous. Furthermore, given **A3** and **A4**, and the boundedness condition on W, it follows that T^*W is also bounded. Therefore, T^*W satisfies (i) of the definition of $\mathcal{M}_b$. To see that T^*W is homogeneous of degree one, let (a^*, γ^*) be a solution to the *saddle-point* Bellman equation at (x, μ, s). Then, for any $\lambda > 0$

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

Since $W(x^{*\prime}, \varphi(\mu, \gamma^*), s') = \varphi(\mu, \gamma^*)\omega(x^{*\prime}, \varphi(\mu, \gamma^*), s')$ with ω homogenoeus of degree zero, the last equation implies that $(a^*, \lambda\gamma^*(x, \mu, s)) \in \Psi_{\lambda(T^*W)}(x, \mu, s)$; however, we need to show that $(a^*, \gamma^*(x, \lambda\mu, s)) \in \Psi_{(T^*W)}(x, \lambda\mu, s)$ for some $\gamma^*(x, \lambda\mu, s)$. Let $\gamma^*(x, \lambda\mu, s) \equiv \lambda\gamma^*(x, \mu, s)$, and for any $\gamma \geq 0$ let $\gamma_\lambda \equiv \gamma\lambda^{-1}$, then for any $a \in A(x, s)$ (resulting in $x' = \ell(x, a, s')$) and $\gamma \geq 0$,

$$\begin{aligned}
& \lambda\mu h_0(x, a^*, s) + \gamma h_1(x, a^*, s) + \beta \mathrm{E} W(x^{*\prime}, \varphi(\lambda\mu, \gamma), s') \\
\equiv\ & \lambda\mu h_0(x, a^*, s) + \lambda\gamma_\lambda h_1(x, a^*, s) + \beta \mathrm{E} W(x^{*\prime}, \varphi(\lambda\mu, \lambda\gamma_\lambda), s') \\
=\ & \lambda[\mu h_0(x, a^*, s) + \gamma_\lambda h_1(x, a^*, s) + \beta \mathrm{E} W(x^{*\prime}, \varphi(\mu, \gamma_\lambda), s')] \\
\geq\ & \lambda[\mu h_0(x, a^*, s) + \gamma^*(x, \mu, s) h_1(x, a^*, s) + \beta \mathrm{E} W(x^{*\prime}, \varphi(\mu, \gamma^*(x, \mu, s)), s')] \\
=\ & \lambda\mu h_0(x, a^*, s) + \gamma^*(x, \lambda\mu, s) h_1(x, a^*, s) + \beta \mathrm{E} W(x^{*\prime}, \varphi(\lambda\mu, \gamma^*(x, \lambda\mu, s)), s') \\
\geq\ & \lambda[\mu h_0(x, a, s) + \gamma^*(x, \mu, s) h_1(x, a, s) + \beta \mathrm{E} W(x', \varphi(\mu, \gamma^*(x, \mu, s)), s')] \\
=\ & \lambda\mu h_0(x, a, s) + \gamma^*(x, \lambda\mu, s) h_1(x, a, s) + \beta \mathrm{E} W(x', \varphi(\lambda\mu, \gamma^*(x, \lambda\mu, s)), s').
\end{aligned}$$

The three equalities follow from the above definitions and the fact that W is homogeneous of degree one in μ, while the two inequalities follow from the fact that $(a^*, \gamma^*(x, \mu, s)) \in \Psi_{(T^*W)}(x, \mu, s)$. This shows that $(a^*, \gamma^*(x, \lambda\mu, s)) \in \Psi_{(T^*W)}(x, \lambda\mu, s)$ and, in fact, the second equality shows that $(T^*W)(x, \lambda\mu, s) = \lambda(T^*W)(x, \mu, s)$.

To show that T^*W is convex, choose arbitrary $\alpha \in (0, 1)$, μ, $\widetilde{\mu}$, in R_+^{l+1} and (x, s). Let $\mu_\alpha \equiv \alpha\mu + (1-\alpha)\widetilde{\mu}$, $(a_\alpha^*, \gamma_\alpha^*) \in \Psi_{(T^*W)}(x, \mu_\alpha, s)$, $x_\alpha^{*\prime} = \ell(x, a_\alpha^*, s')$ and $(a^*, \gamma^*) \in \Psi_{(T^*W)}(x, \mu, s)$, $x^{*\prime} = \ell(x, a^*, s')$

$(\widetilde{a}^*, \widetilde{\gamma}^*) \in \Psi_{(T^*W)}(x, \widetilde{\mu}, s), \widetilde{x}^{*\prime} = \ell(x, \widetilde{a}^*, s')$ and $\widetilde{\gamma}^*_\alpha = \alpha\gamma^* + (1-\alpha)\widetilde{\gamma}^*$, then

$$\begin{aligned}
& (T^*W)(x, \mu_\alpha, s) \\
= \; & \mu_\alpha h_0(x, a^*_\alpha, s) + \gamma^*_\alpha h_1(x, a^*_\alpha, s) + \beta \mathrm{E}\left[W(x^{*\prime}_\alpha, \varphi(\mu_\alpha, \gamma^*_\alpha), s') | \, s\right] \\
\leq \; & \mu_\alpha h_0(x, a^*_\alpha, s) + \widetilde{\gamma}^*_\alpha h_1(x, a^*_\alpha, s) + \beta \mathrm{E}\left[W(x^{*\prime}_\alpha, \varphi(\mu_\alpha, \widetilde{\gamma}^*_\alpha), s') | \, s\right] \\
\leq \; & \mu_\alpha h_0(x, a^*_\alpha, s) + \widetilde{\gamma}^*_\alpha h_1(x, a^*_\alpha, s) + \beta \mathrm{E}\left[\alpha W(x^{*\prime}_\alpha, \varphi(\mu, \gamma^*), s') + (1-\alpha) W(x^{*\prime}_\alpha, \varphi(\widetilde{\mu}, \widetilde{\gamma}^*), s') | \, s\right] \\
= \; & \alpha\left[\mu h_0(x, a^*_\alpha, s) + \gamma^* h_1(x, a^*_\alpha, s) + \beta \mathrm{E} W(x^{*\prime}_\alpha, \varphi(\mu, \gamma^*), s')\right] \\
& + (1-\alpha)\left[\widetilde{\mu} h_0(x, a^*_\alpha, s) + \widetilde{\gamma}^* h_1(x, a^*_\alpha, s) + \beta \mathrm{E} W(x^{*\prime}_\alpha, \varphi(\widetilde{\mu}, \widetilde{\gamma}^*), s')\right] \\
\leq \; & \alpha\left[\mu h_0(x, a^*, s) + \gamma^* h_1(x, a^*, s) + \beta \mathrm{E} W(x^{*\prime}, \varphi(\mu, \gamma^*), s')\right] \\
& + (1-\alpha)\left[\widetilde{\mu} h_0(x, \widetilde{a}^*, s) + \widetilde{\gamma}^* h_1(x, \widetilde{a}^*, s) + \beta \mathrm{E} W(\widetilde{x}^{*\prime}, \varphi(\widetilde{\mu}, \widetilde{\gamma}^*), s')\right] \\
= \; & \alpha (T^*W)(x, \mu, s) + (1-\alpha)(T^*W)(x, \widetilde{\mu}, s),
\end{aligned}$$

where the first inequality follows from the fact that γ^*_α is a minimizer at (x, μ_α, s), the second from the convexity of W and the third from the maximality of a^* and $\widetilde{a}^*$ at (x, μ, s) and $(x, \widetilde{\mu}, s)$ respectively■

Proof of Proposition 2(iii): This is just an application of the *Blackwell's sufficiency conditions for a contraction* (e.g. Stokey et al. (1989) Theorem 3.3.). The following Lemmas 6A - 8A show that T^* satisfies the conditions of the *Contraction Mapping Theorem* and *Blackwell's sufficiency conditions.* ■

Lemma 6A. $\mathcal{M}$ is a non-empty complete metric space (recall that $\mathcal{M}$ denotes either $\mathcal{M}_b$ or $\mathcal{M}_{bc}$)

Proof: It follows from the definition of $\mathcal{M}$ that it is non-empty. Without accounting for the homogeneity property, it follows from standard arguments (see, for example, Stokey, et al. (1989), Theorem 3.1) that every Cauchy sequence $\{W^n\} \in \mathcal{M}$ converges to $W \in \mathcal{M}$ satisfying *i*) and the convexity property *ii*) (and *iii*) if $W \in \mathcal{M}_{bc}$). To see that the homogeneity property is also satisfied, note that for any (x, μ, s) and $\lambda > 0$,

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

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 Given (x, μ, s), let $(\widehat{a}^*, \widehat{\gamma}^*)$ and $(\widetilde{a}^*, \widetilde{\gamma}^*)$ be the solutions to $(T^*\widehat{W})$ and $(T^*\widetilde{W})$, respectively. Then,

$$\begin{aligned}
& (T^*\widehat{W})(x, \mu, s) = \underset{\min_{\gamma \geq 0}, \max_{a \in A(x,s)}}{\mathrm{SP}} \{\mu h_0(x, a, s) + \gamma h_1(x, a, s) + \beta \mathrm{E}\widehat{W}(\ell(x, a, s'), \varphi(\mu, \gamma), s')\} \\
= \; & \mu h_0(x, \widehat{a}^*, s) + \widehat{\gamma}^* h_1(x, \widehat{a}^*, s) + \beta \mathrm{E}\widehat{W}(\ell(x, \widehat{a}^*, s'), \varphi(\mu, \widehat{\gamma}^*), s') \\
\leq \; & \mu h_0(x, \widehat{a}^*, s) + \widetilde{\gamma}^* h_1(x, \widehat{a}^*, s) + \beta \mathrm{E}\widehat{W}(\ell(x, \widehat{a}^*, s'), \varphi(\mu, \widetilde{\gamma}^*), s') \\
\leq \; & \mu h_0(x, \widehat{a}^*, s) + \widetilde{\gamma}^* h_1(x, \widehat{a}^*, s) + \beta \mathrm{E}\widetilde{W}(\ell(x, \widehat{a}^*, s'), \varphi(\mu, \widetilde{\gamma}^*), s') \\
\leq \; & \mu h_0(x, \widetilde{a}^*, s) + \widetilde{\gamma}^* h_1(x, \widetilde{a}^*, s) + \beta \mathrm{E}\widetilde{W}(\ell(x, \widetilde{a}^*, s'), \varphi(\mu, \widetilde{\gamma}^*), s') \\
= \; & \underset{\min_{\gamma \geq 0}, \max_{a \in A(x,s)}}{\mathrm{SP}} \{\mu h_0(x, a, s) + \gamma h_1(x, a, s) + \beta \mathrm{E}\widetilde{W}(\ell(x, a, s'), \varphi(\mu, \gamma), s')\} = (T^*\widetilde{W})(x, \mu, s),
\end{aligned}$$

where the second inequality follows from $\widehat{W} \leq \widetilde{W}$, and the first and the third inequalities from the minimality of $\widehat{\gamma}^*$ and the maximality of $\widetilde{a}^*$ respectively. ■

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

Proof: First, note that $(W + r)(x, \mu, s) = \mu\omega(x, \mu, s) + r$, therefore $\Psi_{W+r}(x, \mu, s) = \Psi_W(x, \mu, s)$. Let $(a^*, \gamma^*) \in \Psi_W(x, \mu, s)$, then

$$\begin{aligned}(T^*(W + r))(x, \mu, s) &= \mu h_0(x, a^*, s) + \gamma^* h_1(x, a^*, s) + \beta \left(\mathrm{E}\left[W(x^{*\prime}, \varphi(\mu, \gamma^*), s') | \, s \right] + r \right) \\ &= (T^*W)(x, \mu, s) + \beta r \blacksquare\end{aligned}$$

REFERENCES

Ábrahám, A., E. Carceles-Poveda, Y. Liu and R. Marimon (2017): "On the optimal design of a Financial Stability Fund," European University Institute.

Ábrahám, A. and N. Pavoni (2005): "The Efficient Allocation of Consumption under Moral Hazard and Hidden Access to the Credit Market," *Journal of the European Economic Association*, vol. 3(2-3), 370-381.

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* 146(3), 1009-1041.

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 U.J. Jermann (2000): "Efficiency, Equilibrium, and Asset Pricing with Risk of Default,"*Econometrica* 68, 775–798.

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

Bertsekas, D. (2009): *Convex Optimization Theory.* Athena Scientific.

Broer, T. (2013): "The Wrong Shape of Insurance? What Cross-Sectional Distributions Tell Us about Models of Consumption Smoothing, " *American Economic Journal: Macroeconomics* 5(4), 107-140.

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

Chien, Y., H. Cole and H. Lustig (2012): "Is the Volatility of the Market Price of Risk Due to Intermittent Portfolio Rebalancing?," *American Economic Review* 102(6), 2859-96.

Cole, H. and F. Kubler (2012): "Recursive Contracts, Lotteries and Weakly Concave Pareto Sets" *Review of Economic Dynamics* 15(4), 475-500.

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.

Debortoli, D. and R. Nunes, (2010). "Fiscal policy under loose commitment," *Journal of Economic Theory*, vol. 145(3), 1005-1032, May.

Epple, D., L. Hansen and W. Roberds (1985): "Quadratic Duopoly Models of Resource Depletion," in *Energy, Foresight, and Strategy,* edited by T. J. Sargent, Resources for the Future.

Faraglia, A. Marcet and A. Scott (2014): "Modelling Long Bonds – The Case of Optimal Fiscal Policy," CEPR DP 9956.

Ferrero, G. and A. Marcet (2005): "Limited Commitment and Temporary Punishment", Universitat Pompeu Fabra Ph. D. thesis (Chapter 2).

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. Perri (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., F. Perri and H. Lustig (2008): "Evaluating Asset Pricing Models with Limited Commitment using Household Consumption Data," *Journal of the European Economic Association* 6(2-3), 715-726.

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.

Levine, P. and D. Currie (1987): "The design of feedback rules in linear stochastic rational expectations models," *Journal of Economic Dynamics and Control*, 11(1), 1-28.

Ljungqvist L. and T. Sargent (2012): *Recursive Macroeconomic Theory* (Third Edition). The MIT Press.

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

Lustig. H., C. Sleet and S. Yeltekin (2008): "Fiscal Hedging with Nominal Assets," *Journal of Monetary Economics* 55, (4), 710-727.

Marcet, A. (2008): "Recursive Contracts," in *The New Palgrave 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, and 2011): "Recursive Contracts, " *European University Institute, ECO 1998 #37 WP , Universitat Pompeu Fabra WP # 337, and EUI-MWP, 2011/03, EUI-ECO, 2011/15, Barcelona GSE wp 552.*

Marcet, A. and A. Scott (2009): "Debt and Deficit Fluctuations and the Structure of Bond Markets", *Journal of Economic Theory,* Vol. 144, N. 2, 473-501.

Marimon, R. and V. Quadrini (2011): "Competition, Human Capital and Income Inequality with Limited Commitment", *Journal of Economic Theory*, 146, 978-1008.

Marimon, R. and J. Werner (2017): "The Envelope Theorem, Euler and Bellman Equations, without Differentiability," mimeo, Univeristy of Minnesota and EUI.

Mele A. (2014): "Repeated Moral Hazard and Recursive Lagrangians," *Journal of Economic Dynamics and Control* 42(C), 69-85.

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

Messner M., N. Pavoni and C. Sleet (2013): "The Dual Approach to Recursive Optimization: Theory and Examples," mimeo, IGIER, Bocconi University.

Phelan, C. and E. Stacchetti (2001):"Sequential Equilibria in a Ramsey Tax Model," *Econometrica* 69, 1491-1518.

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,.

Sargent, T.J.(1987): *Macroeconomic Theory*, (Chapter XV), New York: Academic Press, Second edition.

Sleet, C. and S. Yeltekin (2010): "The Recursive Lagrangian Method: Discrete Time," mimeo, Carengie-Mellon University.

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.*

Takayama, A. (1974): *Mathematical Economics*. The Dryden Press.

Thomas, J. and T. Worral (1988): "Self-Enforcing Wage Contracts," *Review of Economic Studies* 55: 541–554.