RECURSIVE CONTRACTS

Albert Marcet

University College of London, Barcelona GSE and CEPR

Ramon Marimon

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

This version: **January, 2019***

Abstract

We obtain a recursive formulation for a general class of optimization problems with *forward-looking* constraints which often arise in economic 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 continuation problem is obtained with μ_t playing the role 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. We would like to thank Árpád Ábrahám, Fernando Alvarez, Truman Bewley, Charles Brendon, Edward Green, Esther Hauk, David Levine, Robert Lucas, Andreu Mas-Colell, Fabrizio Perri, Edward Prescott, Victor Ríos-Rull, Thomas Sargent, Fabian Schütze, Robert Townsend and, especially, 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 which helped us to generalize our original formulation. Support from MCyT-MEyC of Spain, CIRIT, Generalitat de Catalunya and the hospitality of the Federal Reserve Bank of Minneapolis are acknowledged. Marcet has received financial support from 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 and the Severo Ochoa Programme for Centres of Excellence in R&D (SEV-2015-0563).

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 (2018) describe a large number of applications to macroeconomic models. Under standard assumptions 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 derived from 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 do not satisfy this requirement and they include *forward-looking constraints*, where future actions a_{t+n} also constrain the feasible set of a_t. This occurs, for example, in problems where the principal chooses a contract subject to intertemporal participation constraints (see Example 1 below), and in models of optimal policy under 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 the choice for a_t carries with it an implicit promise about a_{t+n}, therefore contracting parties need to keep track of some additional variables summarizing commitments made in the past about today's choice. The absence of a standard recursive formulation greatly complicates the analysis and numerical solution.

In this paper, we provide an integrated approach for a recursive formulation of a large class of dynamic maximization problems with *forward-looking constraints*. Our interest lies in solving a maximization problem $\mathbf{PP}_\mu$ that depends on certain weights μ. 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's 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; the forward-looking constraints are embedded in the objective function of this continuation problem.

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 **SPFE** 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 time-invariant *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 prove *sufficiency* –that is, solutions to the **SPFE** solve the planning problem of interest $\mathbf{PP}_\mu$, when the value function of **SPFE** is differentiable in μ_t for every (x_t, μ_t, s_t); a property which is satisfied when the solution for a_t in the SPFE is unique. For more general cases (e.g. non-concave problems, or non differentiable value functions, possibly with multiple solutions), we provide an *intertemporal consistency condition* (**ICC**) guaranteeing sufficiency. We show that when **SPFE** has solutions there is one satisfying **ICC**, which is easily obtained in computed solutions. Finally we also provide conditions for the existence of *saddle-point* solutions to **SPFE** and show how standard dynamic programming results – such as the contraction property implying uniqueness of the value function – naturally extend to our **SPFE.**

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 $\mathbf{PP}_\mu$. 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.*

$\mathbf{PP}_{\mu_t}$, with a given initial condition (x_t, s_t), is labelled as 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 signals 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). A difficulty in using this approach to find numerical solutions is that 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. But – under standard assumptions – the continuation problem $\mathbf{PP}_{\mu'}$ has a solution for *any* $\mu' \geq 0$; thus, our approach sidesteps 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, Marimon and Quadrini (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)).

Section 2 provides a basic introduction to our approach. The main body of the theory is in Sections 3 and 4 of this paper, while most proofs are in the Appendix. The relation to the literature and the promised utility approach is discussed in Section 5. Section 6 concludes.

2 Formulating contracts as *recursive saddle-point problems*

In this section, we provide an outline of our approach. We show how dynamic programming methods can be extended to find a recursive formulation for a large class of models with *forward-looking constraints.* We leave the formal results to sections 3 and 4. This section should be self-sufficient for a user of the method.

The class of models under study is characterized as dynamic planning problems ($\mathbf{PP}_\mu$) with a return function 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).

Here ℓ, p, h_0, h_1 are known functions; β, x_0, s_0 and $\mu \equiv (\mu^0, ..., \mu^l) \in R_+^{l+1}$ are known constants or vectors, and $\{s_t\}_{t=0}^{\infty}$ an exogenous stochastic Markov process. We denote as h_i^j 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 $a_t(s^t) \in A \subset R^m$ is a state-contingent action, as usual we take $s^t = (s_0, ..., s_t)$.

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 below). Constraints with $N_j = 0$ cover cases where the planner must take into acccount intertemporal reactions of agents, as in dynamic Ramsey equilibria (see Example 2 below) [2].

Letting $(\boldsymbol{a}^*, \boldsymbol{x}^*\ ) = \{a_t^*, x_t^*\}_{t=0}^{\infty}$ denote a solution of PP_μ at (x_0, s_0), the value of the objective function – parameterized by μ – is given by $V_\mu(x_0, s_0) \equiv E_0 \sum_{j=0}^{l} \sum_{t=0}^{N_j} \beta^t \mu^j h_0^j(x_t^*, a_t^*, s_t)$.

It is without loss of generality that the same function h_0 appears in the objective function and in the constraints (3) [3]. Note also that even though μ could be normalised without affecting the solution (for example, taking $\mu^0 = 1$) the value function V_μ is defined for all $\mu \in R_+^{l+1}$. Both of these features are needed to deliver the continuation problem that suitably characterizes a recursive solution in Proposition 1 below.

Standard dynamic programming considers the following special case of $\mathbf{PP}_\mu$: *i)* forward-looking constraints (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 a standard *Bellman functional equation* holds in that case under very general assumptions[4]. This guarantees the powerful result that the optimal solution to $\mathbf{PP}_\mu$ satisfies $a_t^* = \psi_\mu(x_t^*, s_t)$ for a *time-independent* policy function ψ_μ derived from the Bellman equation. This result is very often used in the literature to characterize and compute solutions to $\mathbf{PP}_\mu$. Furthermore, the solution is time-consistent.

[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 5.

[3]Example 2 below substantiates this claim.

[4]More precisely, the value function satisfies $V_\mu(x, s) = \sup_a \{\mu h_0(x,\, a,\, s) + \beta \ \mathrm{E}\,[V_\mu(x', s') \mid s]\}$ s.t. (2). We denote $\mu h_i(x, a, s) \equiv \sum_{j=0}^{l} \mu^j h_i^j(x, a, s)$.

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-inconsistent.

2.1 An intuitive argument

We now provide an intuitive argument showing how the Lagrangian of (1) can be formulated in recursive form, with respect to the constraints (3). This formulation is very convenient technically and conceptually, since using a standard Lagrangian approach provides the basic framework to derive our recursive formulation and enlightens the key feature of our approach: *forward-looking constraints can be summarized in a co-state vector*, μ. A formal analysis is given in Sections 3.

A Lagrangian of $\mathbf{PP}_\mu$, that incorporates forward-looking constraints can be written as

$$\begin{aligned}\mathcal{L}_\mu(\boldsymbol{a},\boldsymbol{\gamma};x_0,\,s_0) \;=\;& \mathrm{E}_0\left[\sum_{j=0}^{l}\sum_{t=0}^{N_j}\beta^t\;\mu^j h_0^j(x_t,\,a_t,\,s_t)+\right. \qquad (4)\\ & \left.\sum_{t=0}^{\infty}\beta^t\sum_{j=0}^{l}\gamma_t^j\;\mathrm{E}_t\sum_{n=1}^{N_j+1}\beta^n\left(h_0^j(x_{t+n},\,a_{t+n},\,s_{t+n})+h_1^j(x_t,\,a_t,\,s_t)\right)\right],\end{aligned}$$

where γ_t is the Lagrange multiplier associated with (3)[5]. To simplify the exposition, the remaining constraints are imposed separately, hence $\mathcal{L}_\mu$ is defined for $\boldsymbol{a}$ satisfying (2).

Using the law of iterated expectations to eliminate E_t and simple algebra one can show that for each argument $(\boldsymbol{a},\boldsymbol{\gamma})$ we can rewrite $\mathcal{L}_\mu$ as[6]:

$$\mathcal{L}_\mu(\boldsymbol{a},\boldsymbol{\gamma};x_0,\,s_0)=\mathrm{E}_0\left[\sum_{t=0}^{\infty}\beta^t\;\left[\mu_t\,h_0(x_t,\,a_t,\,s_t)+\gamma_t\,h_1(x_t,\,a_t,\,s_t)\right]\right], \qquad (5)$$

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

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

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

Upon inspection of (5)-(6) and (2), it should be 'intuitive' that $\mathcal{L}_\mu$ can yield a recursive structure similar to the programs amenable to dynamic programming; namely, the objective function (5) is a discounted sum with time-invariant return functions $(h_0,\,h_1)$ and past shocks enter into the transition functions (6) and (2), and into the return function at t, only through the 'state variables' (x_t,μ_t). This interpretation relies on the fact that $(\boldsymbol{a},\boldsymbol{\gamma})$ are decision variables of the Lagrangian *and* on the introduction of $\boldsymbol{\mu}\equiv\{\mu_t\}_{t=0}^{\infty}$ as a co-state variable with transition function given by (6). This suggests that, to the extent that solutions of $\mathcal{L}_\mu(\cdot;x_0,\,s_0)$ are solutions to $\mathbf{PP}_\mu$, the solution we seek satisfies $(a_t,\gamma_t)=\psi(x_t,\mu_t,s_t)$ for some time-invariant function ψ.

[5] In fact we should refer to γ_t as a "normalised" multiplier. Strictly speaking the Lagrange multiplier of the $j-th$ constraint (2) at t for a realization s^t is given by $\beta^t\gamma_t(s^t)P(s^t\mid s_0)$, where P is the probability measure of s^t.

[6] See Appendix A for the algebra.

2.2 An alternative functional equation

The intuition in the previous paragraph can not be formalised by appealing to standard dynamic programming. This is because the Bellman equation is shown to hold for dynamic *maximisation* problems, but the above Lagrangian – i.e. (5) subject to (6)-(2) – gives the desired solution to $\mathbf{PP}_\mu$ if we find the *saddle-point* of that Lagrangian. Therefore, to conclude that the solution to $\mathbf{PP}_\mu$ has a recursive formulation including μ_t as a co-state one needs to derive an analogous functional equation for saddle-point problems. The task of this paper is to prove the connection between the *saddle-point* functional equation and the problem of interest $\mathbf{PP}_\mu$.

To this end, we first introduce **notation for saddle-point problems**. Given a function $\mathcal{F} : Y \times Z \longrightarrow R$ we define a *saddle-point of* $\mathcal{F}$ as $(y^*, z^*) \subset Y \times Z$ satisfying

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

The problem of finding such a (y^*, z^*) is called a *saddle-point problem*, which we denote as:

$$\text{SP} \inf_{z \in Z} \sup_{y \in Y} \mathcal{F}(y, z).$$

The set of (potentially multiple) saddle-points (y^*, z^*) that solve this problem is denoted

$$\arg \text{SP} \inf_{z \in Z} \sup_{y \in Y} \mathcal{F}(y, z).$$

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

We now define a functional equation analog to Bellman's that characterises recursively a *saddle-point of* $\mathcal{L}_\mu$, denoted $(\boldsymbol{a}^*, \boldsymbol{\gamma}^*)$. This will be useful because, as is well known, under suitable conditions $\boldsymbol{a}^*$ is then a solution to $\mathbf{PP}_\mu$ and $\boldsymbol{\gamma}^*$ are the Lagrange multipliers of constraints (3).

We show that the saddle-point value function $W : X \times R_+^{l+1} \times S \to R$ defined as $W(x, \mu, s) \equiv \mathcal{L}_\mu(\boldsymbol{a}^*, \boldsymbol{\gamma}^*)$ satisfies the following *saddle-point functional equation:*

$$\textbf{SPFE} \quad W(x, \mu, s) = \text{SP} \inf_{\gamma \geq 0} \sup_{a \in A} \{\mu h_0(x, a, s) + \gamma h_1(x, a, s) + \beta \, \text{E}[W(x', \mu', s') | \, s]\} \tag{8}$$

$$\text{s.t. } x' = \ell(x, a, s'), \; p(x, a, s) \geq 0, \tag{9}$$

$$\text{and } \mu' = \varphi(\mu, \gamma). \tag{10}$$

Given a value function W satisfying this **SPFE** in any possible state $(x, \mu, s) \in X \times R_+^{l+1} \times S$, the corresponding saddle-point policy correspondence (*SP policy correspondence)* is defined as

$$\Psi_W(x, \mu, s) = \arg \text{SP} \inf_{\gamma \geq 0} \sup_{a \in A} \{\mu h_0(x, a, s) + \gamma h_1(x, a, s) + \beta \, \text{E}[W(x', \mu', s') | \, s]\}$$

[7]For clarity, we denote $\inf_{z \in Z} \left[\sup_{y \in Y} \mathcal{F}(y, z)\right]$ a sequential problem where first one finds $\sup_{y \in Y} \mathcal{F}(y, z)$ for each given z and the resulting sup (itself a function of z) is minimized over z. It is well known that the ordering may matter for this sequential problem, that is, it may be that $\arg\inf[\sup \mathcal{F}] \neq \arg\sup[\inf \mathcal{F}]$ and $\inf[\sup \mathcal{F}] \neq \sup[\inf \mathcal{F}]$, and in this case a saddle-point may not exist. We focus on problems where the saddle-point exists, and provide conditions guaranteeing existence (Theorem 3).

subject to (9)-(10)[8].

Note that (8) has three additional features that are not found in the Bellman equation: *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.

As with the Bellman equation the **SPFE** gives the solution we seek. Our approach is to show, first, necessity of **SPFE**, namely that, under standard assumptions (convexity of the constrained set, etc.), a solution to $\mathbf{PP}_{\mu}$, $\{a_t^*\}_{t=0}^{\infty}$, satisfies $(a_t^*,\, \gamma_t^*) \in \Psi_W(x_t^*, \mu_t^*, s_t)$, for some $\gamma_t^* \geq 0$. If, in addition,Ψ_W is single valued, we denote the resulting function by ψ_W, the solution satisfies $(a_t^*, \gamma_t^*) = \psi_W(x_t^*,\, \mu_t^*,\, s_t)$, and we call it a *saddle-point policy function* (*SP policy function*). Furthermore, the value function of $\mathbf{PP}_{\mu}$ satisfies this functional equation, that is $W(x,\, \mu,\, s) = V_{\mu}(x, s)$ satisfies the **SPFE** (Theorem 1).

We also provide a set of general conditions[9] guaranteeing sufficiency of **SPFE,** namely, that if a value function W satisfies (8) for all $(x,\, \mu,\, s)$, and $(\boldsymbol{a}^{**},\, \boldsymbol{\gamma}^{**})$ satisfies $(a_t^{**},\, \gamma_t^{**}) \in \Psi_W(x_t^{**}, \mu_t^{**}, s_t)$ then $\boldsymbol{a}^{**}$ is a solution of $\mathbf{PP}_{\mu}$[10].

In sum, from the user's perspective, what needs to be retained 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 follows the recursion $\mu_{t+1}^{j,*} = \mu_t^{j,*} + \gamma_t^{j,*}$ for $j \leq k$, (i.e. for constraints involving discounted sums with $N_j = \infty$,), and it is the previous multiplier $\mu_{t+1}^{j,*} = \gamma_t^{j,*}$ for $j > k$, (i.e. for constraints involving one future period with $N_j = 0$). One needs to initialize $\mu_0^* = \mu$.

The examples in sections 2.2.2 and 2.2.3 show how this idea can be applied to obtain recursive solutions in problems with forward-looking constraints.

2.2.1 Time-inconsistency and the continuation problem

In programs where the standard Bellman equation applies the program is time-consistent: reoptimization at the new state in future periods is also a continuation solution from the original state. However, as is well known, in the presence of forward-looking constraints (3) the solution may be time-inconsistent: the value of a_1^* for a given realisation of s_1 differs from the value a_0^* that would optimise $\mathbf{PP}_{\mu}$ if initial conditions at $t = 0$ were (x_1^*, s_1)[11].

The key to our approach will be 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 $\{a_t^*, x_t^*\}_{t=1}^{\infty}$. To see this intuitively, expand the above Lagrangian (5):

[8] For an explicit definition of the saddle-point inequalities, see (19) and (20) in Section 3.

[9] As we show in Section 3, the constrained set may not be convex.

[10] We are ignoring, in this informal description, some delicate issues related to the fact that Ψ_W may be empty or it may be a multi-valued correspondence.

[11] More formally, leaving explicit the dependence on initial conditions, let $\left\{a_t^*(x_0, s^t)\right\}_{t=0}^{\infty}$ denote the solution of $\mathbf{PP}_{\mu}$. Then, absent forward-looking constraints, time-consistency holds; in particular, $a_0^*\left(\ell(x_0,\, a_0^*\left(x_0, s_0\right), s_1\right), s_1)) = a_1^*(x_0, s^1)$. With forward-looking constraints this equality may not hold.

$$\begin{aligned}
\mathcal{L}_{\mu}(\boldsymbol{a},\boldsymbol{\gamma};x_0,\,s_0) &= \mathrm{E}_0\left[\sum_{t=0}^{\infty}\beta^t\;\left[\mu_t\,h_0(x_t,\,a_t,\,s_t)+\gamma_t\,h_1(x_t,\,a_t,\,s_t)\right]\right]\\
&= \mathrm{E}_0\left[\mu_0\,h_0(x_0,\,a_0,\,s_0)+\gamma_0\,h_1(x_0,\,a_0,\,s_0)\right.\\
&\quad \left.+\beta\sum_{t=0}^{\infty}\beta^t\;\left[\mu_{t+1}\,h_0(x_{t+1},\,a_{t+1},\,s_{t+1})+\gamma_{t+1}\,h_1(x_{t+1},\,a_{t+1},\,s_{t+1})\right]\right]\\
&= \mu_0\,h_0(x_0,\,a_0,\,s_0)+\gamma_0\,h_1(x_0,\,a_0,\,s_0)+\beta\mathrm{E}_0\left[\mathcal{L}_{\mu_1}(\boldsymbol{a}',\boldsymbol{\gamma}';x_1,\,s_1)\right],
\end{aligned}$$

where $\boldsymbol{a}\equiv\{a_t\}_{t=0}^{\infty}$, and $\boldsymbol{a}'\equiv\{a_t\}_{t=1}^{\infty}$ denotes its continuation, similarly for $\boldsymbol{\gamma}$ and $\boldsymbol{\gamma}'$. That is, if $(a_0^*,\,\gamma_0^*)$ is the first component of a saddle-point of $\mathcal{L}_{\mu}(\cdot;x_0,s_0)$ determining $x_1^*=\ell(x,\,a_0^*,\,s_1)$ and $\mu_1^*=\varphi(\mu,\,\gamma_0^*)$, then the saddle-point of $\mathcal{L}_{\mu_1^*}(\cdot;x_1^*,\,s_1)$ must coincide with $\boldsymbol{a}^{*\prime}$[12]. Furthermore, $\mathcal{L}_{\mu_1^*}(\boldsymbol{a}^{*\prime},{\boldsymbol{\gamma}^*}';x_1^*,\,s_1)$ is the Lagrangian of $\mathbf{PP}_{\mu_1^*}$ at $(x_1^*,\,s_1)$, therefore the solution of $\mathbf{PP}_{\mu_1^*}$ coincides with the 'sup' argument of the saddle-point of $\mathcal{L}_{\mu_1^*}$. A formal argument is given in Proposition 1.

In this sense we can say that in our approach $\mathbf{PP}_{\mu_1^*}$ is a continuation problem. This gives the following characterization of time-inconsistency: in cases when $\mu_1^*\neq\mu$ (i.e. $\gamma_0^*\neq 0$), the solution to $\mathbf{PP}_{\mu}$ at (x_1^*,s_1) is generally time-inconsistent; obviously, these are precisely the cases where forward-looking constraints are binding.

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 the necessity of **SPFE**. Second, it shows one key advantage over the promised utility approach of Abreu, Pearce and Stachetti: the only constraint on the co-state variable is that $\mu_t\in R_+^{l+1}$, under mild standard assumptions the continuation problem $\mathbf{PP}_{\mu_{t+1}^*}$ has a solution for *all* $\mu_{t+1}\in R_+^{l+1}$, as it involves maximising a continuous objective function over a compact set. This sidesteps the complications of having to find the set of feasible promised utilities; we give a more thorough discussion in section 5. Third, $\mathbf{PP}_{\mu_t^*}$ 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 often 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.

2.2.2 Example 1: Risk-sharing with limited enforcement

Consider a model of a partnership with limited enforcement, where several agents can share their individual risks and jointly invest in a project which can only be undertaken jointly. There is a single consumption good and $l+1$ infinitely-lived consumers indexed by $j=0,...,l+1$ with standard preferences $E_0\sum_{t=0}^{\infty}\beta^t\;u(c_t^j)$, where c is individual consumption. Agent j receives a random endowment of consumption good y_t^j at time t, $y_t=(y_t^0,\ldots,y_t^l)$. 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[13]. Production of

[12] This is because if the saddle-point of $\mathcal{L}_{\mu_1^*}$ differed from $(\boldsymbol{a}^*,\boldsymbol{\gamma}^*)$ then the latter would not be a saddle-point of $\mathcal{L}_{\mu}$, since the continuation of $\mathcal{L}_{\mu}$ satisfies all the constraints and has the same objective function as $\mathcal{L}_{\mu_1^*}$.

[13] A common assumption is that the outside option is autarky, where agent j consumes only his endowment from t onwards, $v_j^a(y_t)=\mathrm{E}\left[\sum_{n=0}^{\infty}\beta^n u(y_{t+n}^j)\mid y_t\right]$. It should be noted that one can allow for the outside option to be

the consumption good is $F(k,\theta)$, 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, c_t i_t are chosen given information on (θ^t, y^t).

The planner solves

$$\max_{\{c_t,i_t\}} \quad \mathrm{E}_0 \sum_{t=0}^{\infty} \beta^t \sum_{j=0}^{l} \alpha^j \; u(c_t^j) \tag{11}$$

$$\text{s.t. } k_{t+1} = (1-\delta)k_t + i_t,$$

$$F(k_t, \theta_t) + \sum_{j=0}^{l} y_t^j \geq \sum_{j=0}^{l} c_t^j + i_t, \text{ and}$$

$$\mathrm{E}_t \sum_{n=0}^{\infty} \beta^n \; u(c_{t+n}^j) \geq v_j^a(y_t), \qquad \text{for all } j = 0, ..., l \text{ and } t \geq 0, \tag{12}$$

to find Pareto optimal allocations subject to enforcement constraints (12) and initial conditions (k_0, y_0, θ_0).

It is easy to map this planner's problem into our $\mathbf{PP}_\mu$ formulation if we take $\mu \equiv (\alpha^0, \ldots, \alpha^l) \equiv \alpha$, $s \equiv (\theta, y)$; $x \equiv k$; $a \equiv (i, c)$; $\ell(x,\, a,\, s) \equiv (1-\delta)k + i$; $p(x,\, a,\, s) \equiv F(k,\theta) + \sum_{k=0}^{l} y^k - \left(\sum_{k=0}^{l} c^k + i\right)$; $h_0^j(x,\, a,\, s) \equiv u(c^j)$; $h_1^j(x,\, a,\, s) \equiv u(c^j) - v_j^a(y_t)$, $j = 0, ..., l$.

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

$$\mathcal{L}_\mu(\boldsymbol{a}, \boldsymbol{\gamma}; k_0, y_0, \theta_0) = \mathrm{E}_0 \sum_{t=0}^{\infty} \beta^t \sum_{j=0}^{l} \left[\mu_{t+1}^j \; u(c_t^j) - \gamma_t^j \, v_j^a(y_t)\right],$$

for feasible consumption allocations. In this case all the forward looking constraints have $N_j = \infty$ hence $\mu_{t+1} = \mu_t + \gamma_t$ with initial conditions $\mu_0 = \alpha$.

The **SPFE** takes the form

$$W(k,\, \mu, y, \theta) \; = \mathrm{SP} \inf_{\gamma \geq 0} \sup_{c,i} \left\{ \sum_{j=0}^{l} \left[\mu^{j\prime} u(c^j) - \gamma^j v_j^a(y)\right] + \beta \; \mathrm{E}\left[W(k',\, \mu', y', \theta') \,|\, y, \theta\right] \right\} \tag{13}$$

and $\mu' = \mu + \gamma$,

subject to feasibility constraints. Our results in sections 3 and 4 guarantee that $W(k,\, \mu, y, \theta) = V_\mu(k, y, \theta)$ solves this functional equation and, recalling that ψ_W is the saddle-point that solves the SP problem in the right side of (13), the solution to the problem of interest (11) satisfies

$$\begin{aligned} (\gamma_t^*, c_t^*, i_t^*) &= \psi_W(k_t^*,\, \mu_t^*,\, \theta_t,\, y_t) \text{ and} \\ \mu_{t+1}^* &= \mu_t^* + \gamma_t^*, \end{aligned} \tag{14}$$

endogenous; for example, to exit and enter another partnership contract with some transitional cost, which requires to solve a fixed-point problem between the postulated outside options and the realised contracts (e.g. Cooley *et al.* (2004)).

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

The continuation problem $\mathbf{PP}_{\mu_1^*}$ replaces the objective function of (11) by $\mathrm{E}_1 \sum_{t=0}^{\infty} \beta^t \sum_{j=0}^{l} \mu_1^{j,*} u(c_t^j)$ for $\mu_1^* = \alpha + \gamma_0^*$ and initial conditions (k_1^*, y_1, θ_1), leaving technological and forward-looking constraints unchanged. This means that the solution after period $t = 1$ coincides with the solution of the original problem when the weights α of the agents in the objective function of (11) are replaced by the co-state variables μ_1^*; therefore the variable μ_1^* is all that needs to be remembered from the past at $t = 1$.

A solution to the continuation problem PP_{μ_1} exists generically for any $\mu_1 \in R_+^{l+1}$, therefore we completely sidestep the complication of having to compute the set of feasible continuation promised utilities as would happen with the promised-utility approach – see section 5.

The evolution of the weights μ_t^* determines agents' consumption. Every time that the enforcement constraint for agent j is binding $\gamma_t^{j,*} > 0$ thus, given the optimality condition $\frac{u'(c_t^{j,*})}{u'(c_t^{i,*})} = \frac{\mu_t^{i,*}}{\mu_t^{j,*}}$ the ratio $\frac{c_t^{j,*}}{\sum_{i=0}^{l} c_t^i}$ increases "permanently". This avoids default while optimally smoothing consumption to the extent possible. This ratio will decrease in the future if the forward-looking constraint is binding for other agents.

Various papers in the literature have exploited these features to describe the evolution of consumption in several related setups[14]. Various contributions show how this planner's problem can be decentralised[15].

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

$$\mu_{t+1}^j u'(c_t^j) = \beta \ \mathrm{E}_t \left[\mu_{t+2}^j u'(c_{t+1}^j)(F_{k_{t+1}} + 1 - \delta) \right]. \tag{15}$$

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

The existence of a time-invariant policy function (14) is key in finding numerical solutions guaranteeing that (15), the participation and the feasibility constraints hold. A useful property is that the vector μ_t can be normalized – for example, with $\widehat{\mu}_t^j = \mu_t^j / \sum_{i=0}^{l} \mu_t^i$. In section 4 we provide conditions for the existence of a time-invariant policy function (Theorem 3).

2.2.3 Example 2: A Ramsey problem

We present an abridged version of the optimal taxation problem under incomplete markets studied by Aiyagari at al. (2002). This example serves various purposes: it is an example of one-period forward-looking constraints when $N_j = 0$, it demonstrates that there is no loss of generality in having the same h_0 in the return and constraints, and it shows why we need a weight μ^0 in the first element

[14] Among others, Marcet and Marimon (1992) studied one-sided constraints in a small open economy, Broer (2013) characterizes the stationary distribution of consumption, Abraham and Laczo (2018) characterise analytically the solution.

[15] See, among others, in Alvarez and Jermann (2000), Kehoe and Perri (2002), Krueger, Lustig and Perri (2008) and Abraham and Cárceles-Poveda (2010).

of h_0 in the formulation of $\mathbf{PP}_\mu$. It will be useful also in section 5 to compare our approach with the promised utility approach.

A government must finance exogenous random expenditures g with labor tax rates τ and issuing real riskless bonds b, given initial bonds b_0. A representative consumer maximizes utility $E_0 \sum_{t=0}^{\infty} \beta^t \, [u(c_t) + v(e_t)]$ subject to a budget constraint $c_t + b_{t+1} p_t^b = e_t(1 - \tau_t) + b_t$. Here c is consumption and e is effort (e.g. hours worked), p_t^b is the bond price and τ_t tax rates. Since government bonds b_t are riskless and markets are incomplete. 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, Ponzi games are ruled out.

In a Ramsey equilibrium the government chooses optimal taxes and debt subject to competitive equilibrium and full commitment. Using a familiar argument, one can substitute out bond prices and taxes by equilibrium relationships so that the Ramsey equilibrium can be found by solving

$$\max_{\{c_t, b_t\}} E_0 \sum_{t=0}^{\infty} \beta^t \, [u(c_t) + v(e_t)] \tag{16}$$

$$\text{s.t.} \qquad E_t \left[\beta b_{t+1} u'(c_{t+1})\right] \geq u'(c_t)(b_t - c_t) - e_t v'(e_t) \tag{17}$$

given b_0 and for $e_t = c_t + g_t$.

Unlike example 1, the forward-looking constraint (17) involves one-period ahead expectation; furthermore the objective function is not present in the forward-looking constraints. Formally, this problem is a special case of $\mathbf{PP}_\mu$ for variables $s \equiv g$; $x \equiv b$, $a \equiv (c, b')$. Taking $h_0^0(x, a, s') \equiv u(c) + v(e)$ and $\mu = (1, 0)$ ensures that the objective function of $\mathbf{PP}_\mu$ coincides with (16). Letting $\ell(x, a, s') \equiv b'$, $h_0^1(x, a, s') \equiv bu'(c)$, $h_1^1(x, a, s') \equiv u'(c)(c - b) + ev'(e)$ and $N_1 = 0$ makes (17) a special case of (3) for $j = 1 = l$. We can incorporate the objective function h_0^0 as part of a constraint by introducing h_1^0 arbitrarily large, ensuring that $\gamma_t^0 = 0$ so that $\mu_t^0 = 1$ for all t.

The objective function of the Lagrangian (5) becomes

$$\mathcal{L}_\mu(\boldsymbol{a}, \boldsymbol{\gamma}; x, s) = \mathrm{E}_0 \sum_{t=0}^{\infty} \beta^t \, \left[\mu_t^0 \left(u(c_t) + v(e_t)\right) + \mu_t^1 b_t u'(c_t) + \gamma_t^1 \left[u'(c_t)(c_t - b_t) + e_t v'(e_t)\right]\right]. \tag{18}$$

The **SPFE** takes the form

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

The continuation problem $\mathbf{PP}_{\mu_1^*}$ is obtained by replacing the objective function in (16) with $E_1 \sum_{t=0}^{\infty} \beta^t \, [u(c_t) + v(e_t)] + \mu_1^{*,1} b_1^* u'(c_0)$ where $\mu_1^{*,1} = \gamma_0^{*,1}$ and for initial conditions (b_1^*, g_1). In this example $\mu_t^0 = 1$ for all t. Keeping an arbitrary value μ_t^0 guarantees homogeneity of degree zero of the optimal choice with respect to μ, a property that we use in some of our theorems.

The key to finding numerical solutions to this problem is that the optimal policy satisfies $(c_t^*, b_t^*, \gamma_t^*) = \psi_W(b_t^*, \gamma_{t-1}^*, g_t)$ with initial conditions $(b_0^*, \gamma_{-1}^*) = (b_0, 0)$ for a time-invariant ψ_W that satisfies (17) and optimality conditions of the Ramsey problem.

Aiyagari et al (2002) discuss how a near-unit root behavior of γ_t influences optimal debt and taxes and that debt acts as a buffer stock for adverse shocks. Faraglia, Marcet, Oikonomou, Scott (2016) show that the role of the co-state γ_t is to enforce a promised tax cut that, in equilibrium, lowers current interest rate costs for a government currently facing high deficits. Various papers exploit and extend the recursive formulation described here in models of Ramsey taxation[16].

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

This section contains the main result of this paper, namely, that the maximization problem $\mathbf{PP}_\mu$ is equivalent to the **SPFE**, under fairly general conditions. In particular, we show *necessity*: solutions to $\mathbf{PP}_\mu$ are solutions to the *saddle-point functional equation* **SPFE** (Theorem 1), we also show that $\mathbf{PP}_{\mu_1^*}$ defines the continuation problem in our approach (proposition 1), this formalizes the discussion in section 2.2.1. We also show *sufficiency*: if the **SPFE** value function W is differentiable in μ, then under minimal additional assumptions, its allocation-solution is a solution to $\mathbf{PP}_\mu$ (Theorem 2). We close the section showing that if the allocation-solution is unique then W is differentiable in μ (Lemma 1) and that in the absence of differentiability a more general *Intertemporal Consistency Condition*, which is always satisfied when a solution to **SPFE** exists (Corollary to Theorem 2), ensures sufficiency. First, we lay out the different assumptions used to obtain these results.

3.1 Assumptions about $\mathbf{PP}_\mu$

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 continuous on (x, a) and, given (x, a), they are $\mathcal{S}$ measurable.

A3. 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 $t \geq 0$.

A4. The functions $h_i^j : X \times A \times S \to R$, $i = 0, 1$, $j = 0, ..., l$, are uniformly bounded, continuous on (x, a) and, given (x, a), they are $\mathcal{S}$ measurable. 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.

[16]Among others, Faraglia, Marcet, Oikonomou and Scott (2019) in a model where the government has to choose a portfolio of maturities; Marcet and Scott (2009) in a model with capital; Schmitt-Grohé, and Uribe (2004) and Siu (2004) introduce nominal bonds and the role of monetary policy; Adam and Billi (2006) introduce a zero-lower bound to interest rates.

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 $(x,\,s)$, and $j=,0,\ldots,l$, there exists a program[17] $\{\widetilde{a}_t\}_{t=0}^{\infty}$, with initial conditions (x,s), satisfying (2), such that $\mathrm{E}_0\sum_{t=1}^{N_j+1}\beta^t h_0^j(\widetilde{x}_t,\,\widetilde{a}_t,\,s_t)+h_1^j(x,\,\widetilde{a}_0,\,s)>0$ and, for $i\neq j$, $\mathrm{E}_0\sum_{t=1}^{N_i+1}\beta^n h_0^i(\widetilde{x}_t,\,\widetilde{a}_t,\,s_t)+h_1^i(x,\,\widetilde{a}_0,\,s)\geq 0$.

A7s. In addition to **A7**, there is an $\epsilon>0$ such that for all $(x,\,s)$, and $j=,0,\ldots,l$, the inequality in **A7** can be replaced by $\mathrm{E}_0\sum_{t=1}^{N_j+1}\beta^t h_0^j(\widetilde{x}_t,\,\widetilde{a}_t,\,s_t)+h_1^j(x,\,\widetilde{a}_0,\,s)\geq\epsilon$.

Assumptions **A1-A3,** are standard, they hold in most applications, and we treat them as our basic assumptions. **A4** guarantees bounded returns and does not preclude sustained growth of the endogenous state x (provided its growth rate is lower than β^{-1})[18]. Assumptions **A5-A6** – in particular the concavity of the h_1^j functions[19] – are not satisfied in some models of interest, however are not used in our sufficiency results (e.g. Theorem 2). Assuming linearity of ℓ in **A5** is the natural consequence of decomposing the action, or control, a from the endogenous state x – which in many applications allows for a reduction of the dimension of the state space, – while keeping convexity of the overall feasibility set[20]. **A7** is a standard interiority assumption (the Slater condition), only needed to guarantee the existence of Lagrange multipliers in R_+^{l+1} that guarantees the saddle point and **A7s** guarantees that the sequence of multipliers is uniformly bounded (Theorem 3)[21].

3.2 The recursive formulation of $\mathbf{PP}_\mu$ (Necessity)

We first show that, under certain standard assumptions, solutions to $\mathbf{PP}_\mu$ satisfy **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*) $\Psi: X\times R_+^{l+1}\times S\rightarrow A\times R_+^{l+1}$ (i.e. $\Psi_W(x,\mu,s)$ is a subset of $A\times R_+^{l+1}$). is:

$$\begin{aligned}
\Psi_W(x,\,\mu,\,s) &= \{(a^*,\gamma^*)\in A\times R_+^{l+1} \text{ satisfying } p(x,a^*,\,s)\geq 0 \text{ s.t.}\\
&\quad \mu h_0(x,\,a^*,\,s)+\gamma h_1(x,\,a^*,\,s)+\beta\mathrm{E}\left[W(\ell(x,\,a^*,\,s'),\,\varphi(\mu,\,\gamma),\,s')|\,s\right]\\
&\geq \mu h_0(x,\,a^*,\,s)+\gamma^* h_1(x,\,a^*,\,s)+\beta\mathrm{E}\left[W(\ell(x,\,a^*,\,s'),\,\varphi(\mu,\,\gamma^*),\,s')|\,s\right] \qquad (19)\\
&\geq \mu h_0(x,\,a,\,s)+\gamma^* h_1(x,\,a,\,s)+\beta\,\mathrm{E}\left[W(\ell(x,\,a,\,s'),\,\varphi(\mu,\,\gamma^*),\,s')|\,s\right] \qquad (20)\\
&\quad \text{for all } \quad (a,\gamma)\in A\times R_+^{l+1} \text{ satisfying } p(x,a,\,s)\geq 0\}.
\end{aligned}$$

[17] We will refer to it as the j-interior program.

[18] Our theory can be extended to unbounded returns in the same way that standard dynamic programming can (see, for example, Alvarez and Stokey (1998)). For simplicity, we focus here on the case of bounded returns.

[19] Note, however, that this assumption can be relaxed since what is needed is the convexity of the constraint set (3).

[20] More precisely, convexity of $\Gamma(\cdot,s,s')$, where $\Gamma(x,s,s')=\{x':\exists\, a\in A \text{ s.t.}\, p(x,\,a,\,s)\geq 0 \text{ and } x'=\ell(x,\,a,\,s')\}$ (e.g. Stokey *et al.* (1989) Ass. 4.8).

[21] One can show that for any (x,s) there exists a solution to $\mathbf{PP}_\mu$ if **A1-A6** are satisfied (Proposition 1 in the 2011 version of this paper).

The results below assume existence of a saddle point $(\boldsymbol{a}^*, \gamma_0^*)$. The reader can think of this as a saddle point of the Lagrangian (5) with $\gamma_0^* \in R^{l+1}$ being the Lagrange multiplier of the *forard-looking* constraint in period $t = 0$. The formal definition of the saddle-point problem that we use, namely $\mathbf{SPP}_\mu$, is in Appendix B.

The following theorem guarantees that the value function V_μ and the solution of $\mathbf{PP}_\mu$ satisfy **SPFE.**

Theorem 1 ($\mathbf{PP}_\mu \Rightarrow$ SPFE). Assume **A1-A4.** Assume, for any $\mu \in R_+^{l+1}$ and any initial condition (x, s), there is a saddle-point $(\boldsymbol{a}^*, \gamma_0^*)$ of $\mathbf{SPP}_\mu$. Then $\boldsymbol{a}^*$ solves $\mathbf{PP}_\mu$, the function $W(x, \mu, s) \equiv V_\mu(x, s)$ satisfies the **SPFE** (8) and $(a_0^*, \gamma_0^*) \in \Psi_W(x, \mu, s)$.

Proof: See Appendix B.

This result assumes the existence of a saddle-point $(\boldsymbol{a}^*, \gamma_0^*)$ of the Lagrangean in $\mathbf{SPP}_\mu$. This assumption is a standard way to proceed in optimization theory, see for example section 8.4, Luenberger (1969). Existence can be checked directly in a given model for a solution obtained using a number of algorithms at hand that can solve $\mathbf{SPP}_\mu$ using our recursive formulation.

The existence of a saddle-point $(\boldsymbol{a}^*, \gamma_0^*)$ can be guaranteed if we strengthen the assumptions of Theorem 1 by requiring concavity and interiority – formally:

Corollary to Theorem 1. 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). The function $W(x, \mu, s) = V_\mu(x, s)$ satisfies the **SPFE** (8) and there is a $\gamma_0^* \in R_+^l$ such that $(a_0^*, \gamma_0^*) \in \Psi_W(x, \mu, s)$.

Proof: See Appendix B.

Note that the results of the Corollary can be obtained from assumptions on the primitives. However, Theorem 1 holds more generally – for example, there are many problems where the feasible set of $\mathbf{PP}_\mu$ is not convex but its solution $\boldsymbol{a}^*$ has a saddle-point $(\boldsymbol{a}^*, \gamma_0^*)$ of $\mathbf{SPP}_\mu$ (e.g. in Example 2, functions h_0^1, h_1^1 may not satisfy **A6**, nevertheless Theorem 1 applies).

The following result shows that $\mathbf{PP}_{\mu_1^*}$ is the appropriate continuation problem in our approach, formalizing our discussion in Subection 2.2.1.

Proposition 1 (Continuation Problem): Assume **A1-A4.** Fix $\mu \in R_+^{l+1}$. Assume that $\mathbf{SPP}_\mu$ has a saddle-point $(\boldsymbol{a}^*, \gamma_0^*)$, hence $\boldsymbol{a}^*$ solves $\mathbf{PP}_\mu$. Then, the continuation of this solution, namely $\{a_t^*\}_{t=1}^\infty$, solves $\mathbf{PP}_{\mu_1^*}$ at (x_1^*, s_1) almost surely in s_1, where $x_1^* = \ell(x, a_0^*, s_1)$ and $\mu_1^* = \varphi(\mu, \gamma_0^*)$.

Proof: See Appendix B.

Note that if $\boldsymbol{a}^*$ solves $\mathbf{PP}_\mu$ at (x, s) and $\mu_1^* \neq \mu$ then the solution of $\mathbf{PP}_\mu$ at (x_1^*, s_1) may differ from the continuation of $\boldsymbol{a}^*$. As explained in Subection 2.2.1 in this case there is time-inconsistency[22]. The results in this section guarantee that even under time-inconsistency the solution can be formulated recursively using the co-state μ_t^*.

A result analogous to the above Corollary can be stated as follows: if assumptions **A5-A6** and **A7** are also required, then the continuation of any solution to $\mathbf{PP}_\mu$ solves $\mathbf{PP}_{\mu_1^*}$.

[22]Strictly speaking time-inconsistency arises generically if there is no scalr ξ such that $\mu = \xi\mu_1^*$.

3.3 The sufficiency of SPFE

We now turn to our sufficiency theorem: **SPFE** $\Rightarrow$ $\mathbf{PP}_\mu$, where the value function W, satisfying the **SPFE** (8), is assumed to be continuous in (x, μ) and convex and homogeneous of degree one in μ, for every s, properties which are satisfied by the Lagrangian $\mathcal{L}_\mu$ – as a function of (x, μ) – associated with the value function V_μ of $\mathbf{PP}_\mu$. We obtain this result assuming that W is also differentiable in μ, a property that is satisfied when the solution $\mathbf{a}^*$ generated by Ψ_W is unique (Lemma 1). In the next subsection we dispense with this assumption and replace it with a weaker *intertemporal consistency condition* (**ICC**), which is satisfied when W is differentiable in μ: the intertemporal Euler equation with respect to μ must be satisfied. We also show that when **SPFE** has a solution – possibly, not unique – there is always a solution satisfying **ICC** (Corollary to Theorem 2).

Theorem 2 (SPFE $\Rightarrow$ $\mathbf{PP}_\mu$). Assume **A4** and that W, satisfying the **SPFE**, is continuous in (x, μ) and convex, homogeneous of degree one and differentiable 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 $\lim_{t\to\infty} \beta^t W(x_t^*,\, \mu_t^*,\, s_t) = 0$, then $\mathbf{a}^*$ is a solution to $\mathbf{PP}_\mu$ at (x, s), and $V_\mu(x,\, s) = W(x,\, \mu,\, s)$.

As a theorem of sufficiency the main assumption is the existence of a *saddle-point Bellman equation* (**SPFE**) with its corresponding solution, but the assumptions on the h_i^j functions are minimal – in particular, we assume boundedness (**A4**) but not concavity – and with respect to W the only 'stringent' assumption is its differentiability with respect to μ. An assumption which – as Lemma 1 shows – is satisfied if the solution $\mathbf{a}^*$ is unique, as it is the case when W is concave and the h_0^j functions strictly concave in x (i.e. **A6s**)[23].

Proof of Theorem 2: The proof is divided into two parts. Part I shows that when W satisfies **SPFE** (8) then the *forward looking constraints* of $\mathbf{PP}_\mu$ are satisfied and W takes the form of the objective function of $\mathbf{PP}_\mu$. Part II shows that $\mathbf{a}^*$ is a maximal element of $\mathbf{PP}_\mu$ and, therefore, that $V_\mu(x,\, s) = W(x,\, \mu,\, s)$ (see Appendix B). The differentiability assumption is only used in Part I.

Part I: Note that if W is homogeneous of degree one and differentiable in μ then, by Euler's Theorem, it has a unique representation $W(x,\, \mu,\, s) = \sum_{j=0}^{l} \mu^j \omega^j(x,\, \mu,\, s)$, where ω^j is the partial derivative of W with respect to μ^j. Given this *Euler representation* the minimality condition (19) takes the form

$$\begin{aligned} & \mu h_0(x,\, a^*,\, s) + \gamma h_1(x,\, a^*,\, s) + \beta \mathrm{E}\left[\varphi(\mu,\, \gamma)\omega(\ell(x,\, a^*,\, s'),\, \varphi(\mu,\, \gamma^*),\, s')|\, s\right] \\ \geq \quad & \mu h_0(x,\, a^*,\, s) + \gamma^* h_1(x,\, a^*,\, s) + \beta \mathrm{E}\left[\varphi(\mu,\, \gamma^*)\omega(\ell(x,\, a^*,\, s'),\, \varphi(\mu,\, \gamma^*),\, s')|\, s\right], \end{aligned} \tag{21}$$

and, by convexity of W in μ, it is satisfied, if and only if, the following *Kuhn-Tucker* conditions are satisfied[24]:

$$h_1^j(x,\, a^*,\, s) + \beta\ \mathrm{E}\left[\omega^j(x^{*\prime},\, \varphi(\mu,\, \gamma^*),\, s')|\, s\right] \quad \geq \quad 0, \tag{22}$$

$$\gamma^{*j}\left[h_1^j(x,\, a^*,\, s) + \beta\ \mathrm{E}\left[\omega^j(x^{*\prime},\, \varphi(\mu,\, \gamma^*),\, s')|\, s\right]\right] \quad = \quad 0. \tag{23}$$

[23] See, for example, Theorem 4.8 in Stokey *et al.* (1989).

[24] Note that in the left-hand side of (21) we have $\omega(\ell(x,\, a^*,\, s'),\, \varphi(\mu,\, \gamma^*),\, s')$ instead of $\omega(\ell(x,\, a^*,\, s'),\, \varphi(\mu,\, \gamma),\, s')$. This follows from the fact that (21) and (19) have the same Kuhn-Tucker conditions (22) and (23); see F4. Lemma 4A in Appendix C.

Alternatively, in order to obtain a Euler equation for the intertemporal minimization problem, the minimality condition (21) can also be written as a choice of μ': for $j = 0, ...k$, $\mu'^j \geq \mu^j$ (i.e. $\mu'^j - \mu^j = \gamma^j \geq 0$) and for $j = k+1, ...l$, $\mu'^j \geq 0$ (i.e. $\mu'^j = \gamma^j \geq 0$), in which case the Envelope Theorem, with respect to μ, takes the form:

$$\partial_{\mu^j} W(x^*,\, \mu^*,\, s) = \omega^j(x^*,\, \mu^*,\, s) = \begin{cases} h_0^j(x^*,\, a^*,\, s) - h_1^j(x^*,\, a^*,\, s) + \lambda^{j*} & \text{if } j = 0, ...k, \\ h_0^j(x^*,\, a^*,\, s) & \text{if } j = k+1, ..., l, \end{cases} \tag{24}$$

where λ^{j*} is the Lagrange multiplier for the constraint $\mu'^{j*} - \mu^{j*} \geq 0$. Therefore, for $j = k+1, ..., l$, $\omega^j(x^*,\, \mu^*,\, s)$ is already defined and, for $j = 0, ...k$, we use the first-order condition with respect to μ'^j, to obtain:

$$h_1^j(x^*,\, a^*,\, s) + \beta \ \mathrm{E}\left[\omega^j(x'^*,\, \mu'^*,\, s_1'|\, s\right] - \lambda^{j*} = 0. \tag{25}$$

Substituting (25) into (24) results in:

$$\omega^j(x^*, \mu^*, s) = \begin{cases} h_0^j(x^*,\, a^*,\, s) + \beta \mathrm{E}[\omega^j(x'^*, \mu'^*, s')|s] & \text{if } j = 0, ...k, \\ h_0^j(x^*,\, a^*,\, s) & \text{if } j = k+1, ..., l, \end{cases} \tag{26}$$

Note that, for $j = 0, ...k$, the equation is the intertemporal Euler equation – that is, in our approach it is a result of the dynamic optimization problem, while in the 'promised-utility' approach it is a constraint: the 'promise-keeping' constraint.

The boundedness assumption **A4**, together with (22) and $\lim_{t\to\infty} \beta^t W(x_t^*,\, \mu_t^*,\, s_t) = 0$, imply that $\lim_{t\to\infty} \beta^t \omega^j(x_t^*,\, \mu_t^*,\, s_t) = 0$, for $j = 0, ...k$. Therefore, we can iterate (26) and obtain:

$$\omega^j(x_t^*,\, \mu_t^*,\, s_t) = \mathrm{E}_t \sum_{n=0}^{N_j} \beta^n h_0^j(x_{t+n}^*,\, a_{t+n}^*,\, s_{t+n}). \tag{27}$$

Equation (27) has two implications. First, it shows that the *Kuhn-Tucker* conditions (22) can be expressed as:

$$h_1^j(x_t^*,\, a_t^*,\, s_t) + \beta \mathrm{E} \sum_{n=0}^{N_j} \left[\beta^n h_0^j(x_{t+n+1}^*,\, a_{t+n+1}^*,\, s_{t+n+1})|\, s_t\right] \geq 0, \quad \text{for} \quad j = 0, ..., l \text{ and } t \geq 0;$$

in other words, that when W is differentiable in μ, solutions to **SPFE** satisfy the *forward-looking* constraints of $\mathbf{PP}_\mu$. Second, it shows that the unique *Euler representation* of W at $(x,\, \mu,\, s)$ is

$$W(x,\, \mu,\, s) = \sum_{j=0}^{l} \mu^j \omega^j(x,\, \mu,\, s) = \mathrm{E}\left[\sum_{j=0}^{l} \sum_{t=0}^{N_j} \beta^n \mu^j h_0^j(x_t^*,\, a_t^*,\, s_t)|s\right], \tag{28}$$

with $(x_0^*, s_0) = (x, s)$. That is, W takes the form of the objective function of $\mathbf{PP}_\mu$. These are the two results we wanted to obtain in Part I ■

Uniqueness and sufficiency without differentiability

If W satisfies **SPFE,** for any (x, s), the function $W\ (x, \cdot,\ s) : R_+^{l+1} \to R$ is finite and, we assume, it is continuous and convex, therefore it is *almost surely* differentiable – i.e. for almost any $\mu \in R_+^{l+1}$ it is differentiable (Rockafellar (1970) Theorem 25.5). However, W is an endogenous function and, in particular, at $(x_t^*,\ \mu_t^*,\ s_t)$ the value function W may be non-differentiable with probability one since $(x_t^*,\ \mu_t^*)$ is an endogenous choice; in other words, while non-differentiability with respect to μ may not be an issue 'at the start' it can be a problem 'along a solution path'. Furthermore, differentiability of W may not be an easy condition to check. To analyse these issues and to obtain sufficiency results (**SPFE** $\Rightarrow$ $\mathbf{PP}_\mu$), when W is not necessarily differntiable, we use subdifferential calculus[25].

Let $\partial_\mu W(x,\ \mu,\ s)$ denote the *subdifferential* of W at $(x,\ \mu,\ s)$ with respect to μ – i.e.

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

For any $\omega(x,\ \mu,\ s) \in \partial_\mu W(x,\ \mu,\ s)$, W has a *Euler representation* $W(x,\ \mu,\ s) = \mu\omega(x,\ \mu,\ s)$. We call $\omega(x,\ \mu,\ s)$ an *Euler representation selection.* In particular, if $\omega_t(x_t^*,\ \mu_t^*,\ s_t) \in \partial_\mu W(x_t^*,\ \mu_t^*,\ s_t)$ there are selections $\omega_t(x_{t+1}^*,\ \varphi(\mu_t^*,\ \gamma_t^*),\ s_{t+1}) \in \partial_\mu W(x_{t+1}^*,\ \varphi(\mu_t^*,\ \gamma_t^*),\ s_{t+1})$ – for every s_{t+1}, following s_t– satisfying:

$$\begin{aligned}
& \mu\omega_t(x_t^*,\ \mu_t^*,\ s_t) = \\
& \mu h_0(x_t^*,\ a_t^*,\ s_t) + \gamma_t^* h_1(x_t^*,\ a_t^*,\ s_t) + \beta \mathrm{E}\left[\varphi(\mu_t^*,\ \gamma_t^*)\omega_t(\ell(x_t^*,\ a_t^*,\ s_{t+1}),\ \varphi(\mu_t^*,\ \gamma_t^*),\ s_{t+1}) |\, s_t\right] \\
\leq\ & \mu h_0(x_t^*,\ a_t^*,\ s_t) + \gamma h_1(x_t^*,\ a_t^*,\ s_t) + \beta \mathrm{E}\left[\varphi(\mu_t^*,\ \gamma)\omega_t(\ell(x_t^*,\ a_t^*,\ s_{t+1}),\ \varphi(\mu_t^*,\ \gamma_t^*),\ s_{t+1}) |\, s_t\right], \qquad (29)
\end{aligned}$$

for all $\gamma \in R_+^{l+1}$, and a^* is a maximal element in the corresponding saddle-point problem (i.e. given the selections and γ^*). Furthermore, the corresponding Kuhn-Tucker (complementary slackness) conditions

$$h_1^j(x_t^*,\ a_t^*,\ s_t) + \beta\ \mathrm{E}\left[\omega_t^j(x_{t+1}^*, \mu_{t+1}^*, s_{t+1}) |\, s_t\right] \geq 0, \qquad (30)$$

$$\gamma_t^{*j}\left[h_1^j(x,\ a_t^*,\ s_t) + \beta\ \mathrm{E}\left[\omega_t^j(x_{t+1}^*, \mu_{t+1}^*, s_{t+1}) |\, s_t\right]\right] = 0, \qquad (31)$$

are necessary and sufficient for (29) to be satisfied (Lemma 5A in Appendix C).

The subindex t in $\omega_t(x_t^*,\ \mu_t^*,\ s_t) \in \partial_\mu W(x_t^*,\ \mu_t^*,\ s_t)$ denotes that the *Euler representation selection* is made at $(x_t^*,\ \mu_t^*,\ s_t)$ and $\omega_t(x_{t+1}^*,\ \varphi(\mu_t^*,\ \gamma_t^*),\ s_{t+1})$ denotes a contingent selection of $\partial_\mu W(x_{t+1}^*,\ \varphi(\mu_t^*,\ \gamma_t^*),\ s_{t+1})$ made at $(x_t^*,\ \mu_t^*,\ s_t)$, while choosing a_t^*.

The value $W(x,\ \mu,\ s)$ is independent of its *Euler representations*; in particular, $\mu_{t+1}^*\omega_{t+1}(x_{t+1}^*, \mu_{t+1}^*, s_{t+1}) = \mu_{t+1}^*\omega_t(x_{t+1}^*,\ \mu_{t+1}^*,\ s_{t+1})$. However, for $j = 0, ..., k$, it may be the case that $\omega_{t+1}^j(x_{t+1}^*, \mu_{t+1}^*, s_{t+1}) \neq \omega_t^j(x_{t+1}^*, \mu_{t+1}^*, s_{t+1})$; in other words the selection of $\partial_\mu W(x_{t+1}^*, \mu_{t+1}^*,\ s_{t+1})$ made at $(x_{t+1}^*, \mu_{t+1}^*, s_{t+1})$ may be *inconsistent* with the contingent selection made at $(x_t^*,\ \mu_t^*,\ s_t)$, which can only happen if $\partial_\mu W(x_{t+1}^*,\ \mu_{t+1}^*,\ s_{t+1})$ is not a singleton – i.e. if W is not differentiable, with respect to μ, at μ_{t+1}^* –, resulting in multiple saddle-point solutions. In fact, this *inconsistency* is the problem that may arise when W is not differentiable. For instance, Messner and Pavoni's (2004) example relies on this inconsistency to show that there are cases where solutions to **SPFE** are not solutions to $\mathbf{PP}_\mu$. We now discuss three different conditions guaranteeing that such an inconsistency problem does not arise. However, before we can state these conditions we need to develop more our results.

[25] See Appendix C for definitions and supporting results.

Our starting point is the *Euler representation* (28), which we have derived in the proof of Theorem 2 (Part I) using the Kuhn-Tucker conditions (22) and differentiability (the Envelope Theorem). In fact, we have derived (27) – the key result to show that the *forward-looking constraints* of $\mathbf{PP}_\mu$ are satisfied – to obtain (28). But, as we now show, the latter can be satisfied even when W is not differentiable in μ. To see this, first note that by (31) the value function has the following recursive representation:

$$\begin{aligned}
W(x_t^*,\, \mu_t^*,\, s_t) &= \mu_t^* \omega_t(x_t^*,\, \mu_t^*,\, s_t) \\
&= \mu_t^* h_0(x_t^*,\, a_t^*,\, s_t) + \gamma_t^* h_1(x_t^*,\, a_t^*,\, s_t) + \beta \, \mathrm{E}\left[W(x_{t+1}^*,\, \varphi(\mu_t^*,\, \gamma_t^*),\, s_{t+1}) |\, s_t \right] \\
&= \mu_t^* h_0(x_t^*,\, a_t^*,\, s_t) + \gamma_t^* h_1(x_t^*,\, a_t^*,\, s_t) + \beta \, \mathrm{E}\left[\varphi(\mu_t^*,\, \gamma_t^*) \omega_t(x_{t+1}^*,\, \varphi(\mu_t^*,\, \gamma_t^*),\, s_{t+1}) |\, s_t \right] \\
&= \mu_t^* h_0(x_t^*,\, a_t^*,\, s_t) + \beta \sum_{j=0}^{k} \mu_t^{*j} \mathrm{E}\left[\omega_t^j(x_{t+1}^*,\, \varphi(\mu_t^*,\, \gamma_t^*),\, s_{t+1}) |\, s_t \right] \\
&\quad + \gamma_t^* \left[h_1(x_t^*,\, a_t^*,\, s_t) + \beta \, \mathrm{E} \omega_t(x_{t+1}^*,\, \varphi(\mu_t^*,\, \gamma_t^*),\, s_{t+1}) |\, s_t \right] \\
&= \mu_t^* h_0(x_t^*,\, a_t^*,\, s_t) + \beta \sum_{j=0}^{k} \mu_t^{*j} \mathrm{E}\left[\omega_t^j(x_{t+1}^*,\, \varphi(\mu_t^*,\, \gamma_t^*),\, s_{t+1}) |\, s_t \right]; \qquad (32)
\end{aligned}$$

however, to have (27) a more strict recursive representation is needed (note the change of subindex on the right-hand side ω):

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

To obtain this representation we need to be more explicit about the fact that solutions $(\mathbf{a}^*, \boldsymbol{\gamma}^*)_{(x,\mu,s)}$ generated by Ψ_W are given by *saddle-point policy selections* ψ_W^s of Ψ_W. In particular, among all solutions it is always possible to choose one where the selection is fixed from the beginning: at (x, μ, s). In other words, one needs to make these choices along the solution path, $(a_t^*, \gamma_t^*) = \psi_W^s(x_t^*,\, \mu_t^*,\, s_t)$, where ψ_W^s is the original selection given by $\psi_W^s(x, \mu, s) \in \Psi_W(x, \mu, s)$ and satisfies $\lim_{t\to\infty} \beta^t W(x_t^*,\, \mu_t^*,\, s_t) = 0$. Given this *saddle-point policy selection* we now sequentially unfold the saddle-point value function W, say from $(x_t^*,\, \mu_t^*,\, s_t)$[26]:

[26] To simplify our expressions we introduce a new 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 $j = k+1, ..., l$.

$$
\begin{aligned}
& \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[W(x_{t+1}^*, \varphi(\mu_t^*, \gamma_t^*), s_{t+1}) | \, s_t\right] \\
= \; & \mu_t^* h_0(x_t^*, a_t^*, s_t) + \gamma_t^* h_1(x_t^*, a_t^*, s_t) \\
& + \beta \, \mathrm{E}\left[\mu_{t+1}^* h_0(x_{t+1}^*, a_{t+1}^*, s_{t+1}) + \gamma_{t+1}^* h_1(x_{t+1}^*, a_{t+1}^*, s_{t+1})\right. \\
& \left. + \beta \, \mathrm{E}\left[W(x_{t+2}^*, \varphi(\mu_{t+1}^*, \gamma_{t+1}^*), s_{t+2}) | \, s_{t+1}\right] |s_t\right] \\
= \; & \mu_t^* \left[h_0(x_t^*, a_t^*, s_t) + \beta \mathrm{E}\left[I^k h_0(x_{t+1}^*, a_{t+1}^*, s_{t+1}) | s_t\right]\right] \\
& + \gamma_t^* \left[h_1(x_t^*, a_t^*, s_t) + \beta \mathrm{E}\left[h_0(x_{t+1}^*, a_{t+1}^*, s_{t+1}) | s_t\right]\right] \\
& + \beta \, \mathrm{E}\left[\gamma_{t+1}^* h_1(x_{t+1}^*, a_{t+1}^*, s_{t+1}) + \beta \, \mathrm{E}\left[W(x_{t+2}^*, \varphi(\mu_{t+1}^*, \gamma_{t+1}^*), s_{t+2}) | \, s_{t+1}\right] |s_t\right] \\
= \; & \mu_t^* \left[h_0(x_t^*, a_t^*, s_t) + \beta \mathrm{E}\left[I^k h_0(x_{t+1}^*, a_{t+1}^*, s_{t+1}) + \beta I^k h_0(x_{t+2}^*, a_{t+2}^*, s_{t+2}) | s_t\right]\right] \\
& + \gamma_t^* \left[h_1(x_t^*, a_t^*, s_t) + \beta \mathrm{E}\left[h_0(x_{t+1}^*, a_{t+1}^*, s_{t+1}) + \beta h_0(x_{t+2}^*, a_{t+2}^*, s_{t+2}) | s_{t+1}\right] s_t\right] \\
& + \beta \, \mathrm{E}\left[\gamma_{t+1}^* \left[h_1(x_{t+1}^*, a_{t+1}^*, s_{t+1}) + \beta h_0(x_{t+2}^*, a_{t+2}^*, s_{t+2})\right] s_{t+1} | s_t\right] \\
& + \beta^2 \, \mathrm{E}\left[W(x_{t+2}^*, \varphi(\mu_{t+1}^*, \gamma_{t+1}^*), s_{t+2} | \, s_t\right] \\
& \ldots \\
= \; & \mu_t^* \left[h_0(x_t^*, a_t^*, s_t) + \beta \mathrm{E}\left[I^k \sum_{n=0}^{T} \beta^n h_0(x_{t+1+n}^*, a_{t+1+n}^*, s_{t+1+n}) | s_t\right]\right] \\
& + \gamma_t^* \left[h_1(x_t^*, a_t^*, s_t) + \beta \mathrm{E}\left[h_0(x_{t+1}^*, a_{t+1}^*, s_{t+1}) + \beta I^k \sum_{n=0}^{T-1} \beta^n h_0(x_{t+2+n}^*, a_{t+2+n}^*, s_{t+2+n}) | s_t\right]\right] \\
& + \beta \mathrm{E}\left[\gamma_{t+1}^* \left[h_1(x_{t+1}^*, a_{t+1}^*, s_{t+1}) + \beta \mathrm{E}\left[h_0(x_{t+2}^*, a_{t+2}^*, s_{t+2}) + \beta I^k \sum_{n=0}^{T-2} \beta^n h_0(x_{t+3+n}^*, a_{t+3+n}^*, s_{t+3+n}) | s_{t+1}\right]\right] |s_t\right] \\
& \ldots \\
& + \beta^T \mathrm{E}\left[\gamma_{t+T}^* \left[h_1(x_{t+T}^*, a_{t+T}^*, s_{t+T}) + \beta h_0(x_{t+T+1}^*, a_{t+T+1}^*, s_{t+T+1})\right] s_{t+T} | s_t\right] \\
& + \beta^{T+1} \mathrm{E}\left[W(x_{t+T+1}^*, \varphi(\mu_{t+T}^*, \gamma_{t+T}^*), s_{t+T+1}) | s_{t+T} | \, s_t\right].
\end{aligned}
$$

Note that, by our boundedness assumption (**A4**), the terms in brackets multiplying the Lagrange multipliers converge, as $T \to \infty$; say, for γ_{t+m}^* to:

$$
\left[h_1(x_{t+m}^*, a_{t+m}^*, s_{t+m}) + \beta \mathrm{E}\left[h_0(x_{t+m+1}^*, a_{t+m+1}^*, s_{t+m+1}) + \beta I^k \sum_{n=0}^{\infty} \beta^n h_0(x_{t+m+2+n}^*, a_{t+m+2+n}^*, s_{t+m+2+n}) | s_{t+m}\right]\right].
$$

But given that the *saddle-point policy selection* ψ_W^s is the same in all iterations, the term in the inner bracket is just $\omega_{t+m}(x_{t+m+1}^*, \mu_{t+m+1}^*, s_{t+m+1})$. Therefore, as $T \to \infty$,

$$
\begin{aligned}
\mu_t^* \omega_t(x_t^*, \mu_t^*, s_t) \; = \; & \mathrm{E}\left[\sum_{j=0}^{l} \sum_{n=0}^{N_j} \beta^n \mu_t^{*j} h_0^j(x_{t+n}^*, a_{t+n}^*, s_{t+n}) | s_t\right] \\
& + \gamma_t^* \left[h_1(x_t^*, a_t^*, s_t) + \beta \mathrm{E}\left[\omega_t(x_{t+1}^*, \mu_{t+1}^*, s_{t+1}) | s_t\right]\right] \\
& + \beta \mathrm{E}\left[\gamma_{t+1}^* \left[h_1(x_{t+1}^*, a_{t+1}^*, s_{t+1}) + \beta \mathrm{E}\left[\omega_{t+1}(x_{t+2}^*, \mu_{t+2}^*, s_{t+2}) | s_{t+1}\right]\right] |s_t\right] \\
& \ldots \\
= \; & \mathrm{E}\left[\sum_{j=0}^{l} \sum_{n=0}^{N_j} \beta^n \mu_t^{*j} h_0^j(x_{t+n}^*, a_{t+n}^*, s_{t+n}) | s_t\right],
\end{aligned}
$$

where the last equality follows from the 'slackness condition' (31). In sum, we have obtained (28) and, in particular, that $\omega_{t+1}^{j}(x_{t+1}^{*},\, \mu_{t+1}^{*},\, s_{t+1}) = \omega_{t}^{j}(x_{t+1}^{*}, \mu_{t+1}^{*},\, s_{t+1}) \equiv \omega^{j}(x_{t+1}^{*}, \mu_{t+1}^{*},\, s_{t+1})$. This derivation of (28) has two implications, which correspond to the first two conditions that guarantee that inconsistency problems do not arise.

First, the role of **uniqueness.** If $\mathbf{a}^*$ is unique then, by (28), the *Euler representation* is unique[27]. However, since the *subdifferential* of W is composed of *Euler representation selections*, this means that $\partial_\mu W$ is a singleton and, therefore, more formally:

Lemma 1. If W, satisfying the **SPFE**, is continuous in (x, μ) and convex in μ, for every s, and for $(\mathbf{a}^*, \boldsymbol{\gamma}^*)_{(x,\mu,s)} \in \Psi_W(x,\, \mu,\, s)$ $\mathbf{a}^*$ is unique, then W is differentiable in μ at $(x,\, \mu,\, s)$.

In fact, what Lemma 1 says is that 'uniqueness' is not a new condition with respect to Theorem 2, but a relatively simple condition to check, which guarantees diferentiability.

Second, the role of **fixing the *saddle-point policy selection***. What our derivation of (28) shows is that if, as it is usually done in computations, the *saddle-point policy selection* is the same in the sequential iterations of **SPFE** the *forward-looking constraints* are consistently defined and, therefore, (27) is satisfied[28]. However, if, at $(x_t^*, \mu_t^*, s_t), W$ is not differentiable in μ and **SPFE** is restarted with a different *saddle-point policy selection* – say, $\psi_W^{\tilde{s}}$ – then, for some j, $\tilde{\omega}_t^{j}(x_t^*,\, \mu_t^*,\, s_t) \neq \omega_{t-1}^{j}(x_t^*,\, \mu_t^*,\, s_t)$, and the resulting solution – up to t with ψ_W^{s} and from t with $\psi_W^{\tilde{s}}$ – may not be a solution to $\mathbf{PP}_\mu$ at (x, s).

Therefore, there is a need to provide a condition (our 'third') guaranteeing consistency that can be checked.

ICC. A solution $(\mathbf{a}^*, \boldsymbol{\gamma}^*)_{(x,\mu,s)}$ generated by the *SP policy correspondence* Ψ_W associated with W satisfies the *Intertemporal Consistency Condition* if, for $t \geq 0$ and $j = 0, ..., k$, its *Euler representation selections* satisfy the intertemporal Euler equation (26); that is if:

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

Corollary to Theorem 2. Assume **A4** and that 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 solutions $(\mathbf{a}^*, \boldsymbol{\gamma}^*)_{(x,\mu,s)}$, satisfying $\lim_{t\to\infty} \beta^t W(x_t^*,\, \mu_t^*,\, s_t) = 0$. If a solution also satisfies the **ICC,** then $\mathbf{a}^*$ is a solution to $\mathbf{PP}_\mu$ at (x, s), and $V_\mu(x,\, s) = W(x,\, \mu,\, s)$. Furthermore, there is a solution which satisfies the **ICC**.

The previous discussion provides the proof to this Corollary, since the only missing piece of the proof of Theorem 2, if differentiability of W in μ is not assumed, is the *Euler equation* (27), which is provided by **ICC** and we have also shown how to obtain a solution that satisfies **ICC**, provided that **SPFE** has a solution. Nevertheless, we have not provided a recursive algorithm that guarantees the *Euler equation* (27) is satisfied.

[27]Note that if, in addition, $\boldsymbol{\gamma}^*$ is also unique then there is a unique *saddle-point policy selection* ψ_W^s; i.e. the *saddle-point policy function* ψ_W.

[28]In the derivation of (28), by keeping the same selection, we had, for $j = 0, ..., k$ and $t > 0$, $\omega_{t+1}^{j}(x_{t+1}^*,\, \mu_{t+1}^*,\, s_{t+1}) = \omega_t^{j}(x_{t+1}^*,\, \mu_{t+1}^*,\, s_{t+1})$.

This can be found in Marimon and Werner (2017) who also provide a more comprehensive discussion of the inconsistency issues discussed here, based on their Envelope Theorem without differentiability which, in our context, generalizes (24).

4 Existence of saddle-point value functions

In this section, we address the issue of the existence of value functions satisfying the **SPFE** (Theorem 3(i)). The existence of saddle-points is needed to show that there is a well-defined contraction mapping generalizing the *Contraction Mapping Theorem* to a *Dynamic Saddle-Point Problem* corresponding to the **SPFE** (Theorem 3(iii)).

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

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

and we also define its subspace of concave functions (in x): $\mathcal{M}_{bc} = \{W \in \mathcal{M}_b \text{ and } iii)\ W(\cdot,\,\mu,\,s) \text{ is concave}\}$. Both spaces are normed vector spaces with the norm

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

We show in Appendix D (Lemma 8A) that these are complete metric spaces and, therefore, suitable spaces for the *Contraction Mapping Theorem.* Note that $V_\mu(x,\,s)$, the value of $\mathbf{PP}_\mu$ with initial conditions (x, s), can also be represented as a function $V_\cdot(\cdot,\cdot)$ – at $(x,\,\mu,\,s)$ – which is in $\mathcal{M}_b$ whenever **A2** - **A4** are satisfied, and in $\mathcal{M}_{bc}$ if, in addition, **A5** - **A6** are satisfied (See Lemma 1A in Appendix B.).

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

$$(T^*W)(x, \mu, s) = \text{SP}\min_{\gamma \ge 0}\max_{a} \left\{\mu h_0(x,\,a,\,s) + \gamma h_1(x,\,a,\,s) + \beta\ \text{E}\left[W(x',\,\mu',\,s')|\,s\right]\right\} \tag{33}$$

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

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

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

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

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[29].

As we have seen in Section 3, any $W \in \mathcal{M}$ has a – possibly non unique – *Euler representation* $W(x, \mu, s) = \mu\omega(x, \mu, s)$ (see also Appendix C). Furthermore, with this representation (a^{*}, γ^{*}) is a *saddle-point* of **SPFE** if, and only if, it is a *saddle-point* of the ***Lagrangian***

$$\begin{aligned}\mathcal{L}(a, \gamma; (x, \mu, s)) \quad &= \quad \mu\left[h_0(x,\, a,\, s) + \beta\ \mathrm{E}\left[\sum_{j=0}^{k} \omega^{j}(x',\, \mu',\, s')|\, s\right]\right] + \gamma\left[h_1(x,\, a,\, s) + \beta\ \mathrm{E}\left[\omega(x', \mu',\, s')|\, s\right]\right], \\ \forall\, \gamma \quad &\in \quad \mathcal{R}_{+}^{l+1},\ \mu' = \varphi(\mu,\, \gamma) \text{ and } a \text{ with } \ p(x, a,\, s) \geq 0,\ x' = \ell(x,\, a,\, s').\end{aligned}$$

Note that γ^{*} plays the double role of being a Lagrange multiplier to the *forward-looking constraints* $h_1(x,\, a,\, s)+$ $\beta\ \mathrm{E}\left[\omega(\ell(x,\ a,\, s'),\, \varphi(\mu,\, \gamma),\, s')|\, s\right] \geq 0$ and an argument in the co-state transition $\varphi(\mu,\, \gamma)$. To prove the existence of such a *saddle-point* we decompose these two roles. First, we show that for any $\widehat{\gamma} \in R_{+}^{l+1}$, in $\varphi(\mu,\, \widehat{\gamma})$, there is a *saddle-point* $(a^{*}(\widehat{\gamma}), \gamma^{*}(\widehat{\gamma}))$, then we use a *fixed point* argument to show that there is a γ^{*} satisfying $(a^{*}(\gamma^{*}), \gamma^{*}(\gamma^{*}))$. The former – i.e. the existence of Lagrange multipliers – requires an interiority (or normality) condition, the latter to strengthen such interiority condition to guarantee that Lagrange multipliers are uniformly bounded. These conditions can take the following form:

IC. W, with $W = \mu\omega$, satisfies the *interiority condition* if, for any $(x,\, s) \in X \times S$, $\mu \in R_{+}^{l+1}$, and $j, j = 0, \ldots l$, there exists $\widetilde{a} \in A$, satisfying $p(x,\, \widetilde{a},\, s) \geq 0$, and, $h_1^{j}(x, \widetilde{a}, s) + \beta \mathrm{E}\left[\omega^{j}(\ell(x,\ \widetilde{a},\, s'),\, \mu, s')|\, s\right] > 0$, and, for $i \neq j$, $h_1^{i}(x, \widetilde{a}, s) + \beta \mathrm{E}\left[\omega^{i}(\ell(x,\ \widetilde{a},\, s'),\, \mu, s')|\, s\right] \geq 0$[30].

SIC. W, with $W = \mu\omega$, satisfies the *strict interiority condition* if it satisfies **IC** and there exists an $\varepsilon > 0$ such that, for any $(x, , s) \in X \times S$, $\mu \in R_{+}^{l+1}$ and $j, \ldots, l$ the inequality $h_1^{j}(x, \widetilde{a}, s) +$ $\beta \mathrm{E}\left[\omega^{j}(\ell(x,\ \widetilde{a},\, s'),\, \mu, s')|\, s\right] > 0$, in **IC** can be replaced by $h_1^{j}(x, \widetilde{a}, s) + \beta \mathrm{E}\left[\omega^{j}(\ell(x,\ \widetilde{a},\, s'),\, \mu, s')|\, s\right]$ $\geq \varepsilon$.

The following lemma, which proof is immediate, provides a condition, easy to check, guaranteeing that these interiority conditions are satisfied.

Lemma 2. $W \in \mathcal{M}$, with $W = \mu\omega$, satisfies **IC** (**SIC**) if for all $(x,\, s) \in X \times S$, $\mu \in R_{+}^{l+1}$, and $j = 0, \ldots l$:

$$\mathrm{E}\left[\omega^{j}(\ell(x,\ \widetilde{a}_0,\, s_0),\, \mu, s_1)|\, s_0\right] \geq \mathrm{E}\left[\sum_{t=0}^{N_j} \beta^{t} h_0^{j}(\widetilde{x}_t,\, \widetilde{a}_{t+1},\, s_{t+1})|\, s_0\right],$$

[29] Recall that Q satisfies the Feller property if whenever f is bounded and continuous in 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).

[30] Note that $\omega^{j}(\ell(x,\ \widetilde{a},\, s'),\, \mu, s')$ can be replaced by $\omega^{j}(\ell(x,\ \widetilde{a},\, s'),\, \varphi(\mu,\, \gamma), s')$, for any $\gamma \in R_{+}^{l+1}$.

for any i-interior program[31] $\{\widetilde{a}_t\}_{t=0}^{\infty}$ of Assumption **A7** (**A7**s), $i = 0, \ldots l$. Furthermore, if $W \in \mathcal{M}$ satisfies **IC** (**SIC**) then T^*W also satisfies **IC** (**SIC**).

In other words, it is enough that $W \in \mathcal{W}$ takes the value of the interior programs of **A7** (**A7**s) as a lower bound to satisfy **IC** (**SIC**); e.g. in the Section 2 example with *limited enforcement constraints* **IC** (**SIC**) is satisfied if W guarantees that at any state $(x,\ s)$, weights $\varphi(\mu,\gamma)$, and j there is an interior (ϵ interior) allocation $\tilde{a}$ that allows agent j satisfy its *forward-looking constraint* with strict inequality (epsilon inequality) while maintaining the *forward-looking constraints* of all the other agents. As Lemma 2 shows, given specific functional forms for $\mathbf{PP}_\mu$ it is not difficult to have $W \in \mathcal{M}$ satisfying these interiority conditions. Note that the last statement of Lemma 2 provides a guide to obtaining $W \in \mathcal{M}$ through value function iteration: start with a value function that satisfies the conditions of Lemma 2.

We can now state the main theorem of this section.

Theorem 3. Assume **A1b** and **A2**-**A5** and **SIC**, and **A6** when $\mathcal{M}$ refers to $\mathcal{M}_{bc}$.

i) Let $W \in \mathcal{M}_{bc}$. 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 (19) - (20). 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.

Theorem 3(*i*) provides conditions for the existence of a *saddle-point*; (*ii*) establishes that the **SPFE** mapping is well defined by showing that T^* maps $\mathcal{M}$ onto itself, and finally (*iii*) shows that T^* is a contraction mapping, therefore there is a unique value function $W \in \mathcal{M}$, $W = T^*W$, satisfying **SPFE**. 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.

Theorem 3 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**). As in standard dynamic programming, if $W \in \mathcal{M}_{bc}$ and the strict concavity assumption **A6s** is satisfied, then $(\mathbf{a}^*)_{(x,\mu,s)}$ is uniquely determined. Also as in standard dynamic programming, if these conditions are not satisfied and *saddle-point* solutions are not unique, an **SPFE** solution is a selection from the *saddle-point* correspondence. However, as we have seen in Section 3, when W is not differentiable in μ a new kind of multiplicity arises[32]. Finally,

[31] See Footnote 17.

[32] Note that it differs from the multiplicity in standard dynamic programming problems – i.e. problems without *forward-looking* constraints – in an important aspect: in a standard dynamic problem if at (x_t^*, s_t) there are multiple solutions, once one is 'selected' leading to $(x_{t+1}^*,\ s_{t+1})$, the latter is a 'a sufficient statistic' in order to follow up on a solution path started at $(x_0,\ s_0)$; in contrast, if at $(x_t^*,\ \mu_t^*,\ s_t)$ there are multiple *saddle-point* solutions (due to the

Theorem 3 also shows that *the contraction property* – very practical for computing value functions – also extends to our *saddle-point Bellman equation operator*.

By Theorem 2, if $W = T^*W$ is differentiable in μ and $(\mathbf{a}^*, \boldsymbol{\gamma}^*)_{(x,\mu,s)}$ is generated by $\Psi_W(x, \mu, s)$, then $\mathbf{a}^*$ is a solution to $\mathbf{PP}_\mu$ at (x, s). Unfortunately, the subspace of differentiable functions is not a complete metric space and, therefore, T^* does not necessarily map $\mu-$differentiable functions into $\mu-$differentiable functions. However, we can provide more structure to T^* to guarantee that the generated solutions $(\mathbf{a}^*, \boldsymbol{\gamma}^*)_{(x,\mu,s)}$ satisfy the *Intertemporal Consistency Condition* **ICC**, and for this we define the T^{**} map.

The $T^{**} : \mathcal{M} \longrightarrow \mathcal{M}$ solves the same *saddle-point* problem as the T^* map, i.e.

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

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

but given $W \in \mathcal{M}$, takes a specific *Euler representation* $W = \mu\omega$ to define the *Euler representation* of $T^{**}W$ according to:

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

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

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

Given that T^{**} solves the same problem as T^*, the results of Theorem 3 hold for T^{**} and there is a $W \in \mathcal{M}$ such that $W = T^*W$ but, in addition, $\omega^j = T^{**}\omega^j$, for $j = 0, ..., l$.Note that, even if W is unique, when it is not differentiable in μ it has multiple *Euler representations* and, correspondingly, the T^{**} map generates multiple solutions. Nevertheless, the **ICC** condition is satisfied. In sum, based on our Corollary to Theorem 2, we have the folowing result:

Corollary to Theorem 3. Let $W \in \mathcal{M}$ satisfy $W = T^{**}W$, for a specific Euler representation $W = \mu\omega$, and $(\mathbf{a}^*, \boldsymbol{\gamma}^*)_{(x,\mu,s)}$ be generated by $\Psi_W(x, \mu, s)$, then $\mathbf{a}^*$ is a solution to $\mathbf{PP}_\mu$ at (x, s).

Note that this corollary provides a guide to the user who is uncertain about whether $W \in \mathcal{M}$ is differentiable in μ: use the T^{**} map and to get the $\mathbf{PP}_\mu$ solution, which simply takes the unique *Euler representation* when W is differentiable in μ, i.e. in this case T^{**} does the same as T^*.

5 Related work

Forward-looking constraints are pervasive in dynamic economic models. Early work introducing Lagrange multipliers as co-state variables in models of optimal policy are found in Epple, Hansen and Roberds (1985), Sargent (1987) and Levine and Currie (1987) in linear-quadratic Ramsey problems,

fact that W is not differentiable in μ), once one is 'selected' leading to $(x^{*\prime}, \mu^{*\prime} \, s')$, the latter may not be a 'sufficient statistic' in order to follow up on a solution path started at $(x_0, \mu_0 \, s_0)$.

justified by the observation that past multipliers appear in the first-order conditions of the Ramsey problem. But this is only indicative of a recursive formulation. Our work provides a formal proof that introducing past multipliers as co-states deliver the optimal solution recursively in a general framework.

The promised-utility approach has been widely used in macroeconomics. Some applications are by Kocherlakota (1995) in a model with participation constraints similar to our Example 1, and Cronshaw and Luenberger (1994) in a dynamic game[33]. 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[34].

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

A key difference between the two approaches lies in the fact that they define very different continuation problems. In the promised-utility approach, promised utility ω_t is a decision today for each possible future state, and this defines a state variable tomorrow, making the problem amenable to a standard Bellman equation treatment. This needs the computation of a correspondence for feasible utilities (denoted $\mathcal{C}_\kappa$) that is very hard to compute. However, as we have emphasied in Section 2 the continuation problem in our approach (namely $\mathbf{PP}_{\mu_1^*}$) is guaranteed to have a solution for any $\mu_1^* \in R_+^{l+1}$. This entirely sidesteps any computation of the feasible set of co-state variables.

We now discuss these issues more concretely by formulating a recursive solution to Example 2 in the context of promised utilities. For ease of exposition assume the exogenous shock g_t is i.i.d. and it can take ν possible values $\overline{g}^\kappa$ $\kappa = 1, ..., \nu$ each with probability π^κ. Constraint (17) can be rewritten as

$$b_{t+1}\beta \sum_{\kappa=1}^{\nu} u'(c_{t+1}(\overline{g}^\kappa))\pi^\kappa = u'(c_t)(b_t - c_t) - e_t v'(e_t). \tag{34}$$

Equation (34) is the 'promise-keeping' constraint and $c_{t+1}(\overline{g}^\kappa)$ is the promised consumption in period $t+1$ if $g_{t+1} = \overline{g}^\kappa$ is realized. The key insight of the promised-utility approach is that by including all promised consumptions $(c_{t+1}(\overline{g}^1), ..., c_{t+1}(\overline{g}^\nu))$ in the vector of today's decision variables a_t, equation (34) 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 $\omega_t = c_t(g_t)$ is included as a co-state variable.

But applying the Bellman equation to this reformulated problem without any further constraint would induce the planner to choose a $c_{t+1}(\overline{g}^\kappa)$ that cannot be supported by any taxation scheme in equilibrium, so in this case the Bellman equation does not provide a feasible solution. To avoid this problem, one needs to compute for each κ the correspondence $\mathcal{C}_\kappa : R \to \mathcal{S}$, where $\mathcal{S}$ is a collection of subsets of R_+ such that if $c_{t+1}(\overline{g}^\kappa) \in \mathcal{C}_\kappa(b_{t+1})$ and if $g_{t+1} = \overline{g}^\kappa$ then a continuation tax and allocation process $\{\tau_{t+j}, c_{t+j}, b_{t+j+1},\}_{j=1}^{\infty}$

[33] Ljungqvist and Sargent (2018) provide an excellent introduction and references to most of this recent work.

[34] As we clarify in this paper – for example, in the discussion of Example 2 below – our approach is not the same as the approach of these papers.

exists that is a competitive equilibrium with $c_{t+1} = c_{t+1}(\overline{g}^{\kappa})$ and corresponding inherited government debt b_{t+1}. Since the correspondence $\mathcal{C}_{\kappa}(\cdot)$ is an endogenous object its computation is very complicated. 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[35], this computation often leads to serious numerical difficulties. Most applications in the literature of the promised-utility approach assume there is no dependence on state variables (i.e. b does not influence C_{κ}) and the sets in S are subsets of R.

As we have seen in Section 2, the issue of computing a feasible set for promised consumption is entirely sidestepped in our approach. This is because any γ^{*1}_{t-1} gives a well-defined continuous objective function of $\mathbf{PP}_{\mu^*_t}$, so that this continuation problem always has a solution[36].

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+1}) in Example 2 under our approach at t, while in the promised utility approach there are $\nu + 1$ decision variables $(c_{t+1}(\overline{g}^1), ..., c_{t+1}(\overline{g}^{\nu}), b_{t+1})$ at t.

As is well known the highest computational savings come from a reduction in the dimension of the state vector. In some cases the recursive Lagrangian has much fewer state variables. Consider generalizing Example 2 to the case where the government issues one long bond that matures in M periods and long bonds are not repurchased by the government, as in Faraglia, Marcet, Oikonomou and Scott (2016, 2019). In this case, the bond price depends on the expectation of marginal utility M periods ahead, so that the analog of (34) gives

$$b^M_{t+1}\beta^M \sum_{\widetilde{g}^M \in G^M} u'(c_{t+M}(g_t, \widetilde{g}^M))\widetilde{\pi}(\widetilde{g}^M) = u'(c_t)(b^M_{t-M+1} - c_t) - e_t v'(e_t), \tag{35}$$

where we denote G^i the set of all possible realizations of $(g_{t+1}, ..., g_{t+i})$, and $\widetilde{\pi}^{\kappa}(\widetilde{g}^M)$ the probability of each sequence. Clearly, the co-state includes $\omega_t = \left(c_t, \left[c_{t+i}(g_t, \widetilde{g}^i)\right]^{i=1,...,M-1}_{\widetilde{g}^i \in G^i}\right)$. For a 10-year bond, even if g only takes two possible values so $\nu = 2$, a quarterly version of the model has more than one trillion state variables, since G^i has 2^i elements[37]. By comparison, the Lagrangian approach can be implemented with 'only' $2M + 1 = 81$ state variables $(\gamma_{t-1}, ..., \gamma_{t-M}, b^M_t, ..., b^M_{t-M+1}, g_t)$[38]

There are some additional differences between the two approaches. Initial conditions for the co-state variables in our approach are known from the outset to be $\mu^0_0 = 1$, $\mu^1_0 = 0$, 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. This is because, as pointed out before, the promised-utility approach determines the variable one period ahead, so it needs an ending boundary condition, while our approach starts out from a given initial condition. It is well known that to find c_0 the Pareto frontier has to be downward sloping; otherwise the computations can become very cumbersome.

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

[36]See the discussion following equation (18).

[37]There are ways of reducing this problem, Lustig et al (2008) provide a recursive formulation with long bonds by adding the yield curve as a state variable. 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.

[38]See Faraglia, Marcet, Oikonomou and Scott (2019) Section 3 for details, and Sections 5, 6, 7 for the state variables in several variations of the model..

The co-state variables in our approach often have an economic interpretation. We have already described in Section 2 how the evolution of μ_t^* can unveil the reason for time-inconsistency problems. Also, μ_t^* can be interpreted as time-varying pareto weights in Example 1 and a time-varying deadweight loss of taxation in Example 2.

Early versions of this paper conceded as an advantage of promised-utility that it could be applied to models under moral hazard and incentive constraints. However, 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. Thus the initial advantages of the promised-utility approach seem to have mostly vanished.

6 Concluding remarks

We have shown that a large class of problems with *forward-looking* constraints can be conveniently formalised as a saddle-point problem. This saddle-point problem obeys a *saddle-point functional equation* (**SPFE**) which is analogous to 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 time-invariant 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 when the policies of a benevolent government are not time-consistent.

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, such as capital or debt, are present; the solution (of a planner's or Ramsey problem) is not time-consistent; our co-state variable μ plays a key rule in determining constrained efficiency *wedges*, or contracts need to be decentralised and, therefore, priced. Similarly, extensions are already available, encompassing a wider set of problems than those considered here (moral hazard, endogenous participation constraints, etc.). Our sufficiency result when the value function is differentiable (in μ) – as in the case that the constrained efficient allocation is unique – already covers a wide range of frequently studied economies. We broaden this range to a larger set of economies (e.g. weakly concave with multiple solutions) by providing the *intertemporal consistency condition* (**ICC**) that must be satisfied when there are *forward-looking* constraints – a condition that is always satisfied when the value function is differentiable (in μ). In the more general case, we show how ICC can be guaranteed when the *saddle-point functional equation* has a solution[39]. Finally, we also provide conditions for the existence of solutions to

[39] Cole and Kubler (2012) provide a generalization to the non-uniqueness case for a restricted class of models. Marimon

our *saddle-point functional equation* (**SPFE**) and extend the main results of dynamic programming to our saddle-point formulation.

Authors' Affiliations:

Albert Marcet, University College of London, Barcelona GSE and CEPR

E-mail: a.marcet@ucl.ac.uk

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

APPENDIX

Appendix A. Rearranging the Lagrangian Here we show that $\mathcal{L}_\mu$ as defined in (4) is equivalent to (5). Shifting E_t in the second line of (4) we can rewrite $\mathcal{L}_\mu$ as

$$\begin{aligned}\mathcal{L}_\mu(\boldsymbol{a},\boldsymbol{\gamma};x,\,s) \quad = \quad & \mathrm{E}_0\Bigg[\sum_{j=0}^{l}\sum_{t=0}^{N_j}\beta^t\;\mu^j h_0^j(x_t,\,a_t,\,s_t)+ \\ & \mathrm{E}_t\sum_{t=0}^{\infty}\beta^t\sum_{j=0}^{l}\gamma_t^j\sum_{n=1}^{N_j+1}\beta^n\left(h_0^j(x_{t+n},\,a_{t+n},\,s_{t+n})+h_1^j(x_t,\,a_t,\,s_t)\right)|s_0\Bigg].\end{aligned} \tag{36}$$

This holds because there is one forward-looking constraint (3) for each possible sequence of shocks s^t, hence γ_t^j is a function of s^t and can be included inside E_t. Using the law of iterated expectations this implies that E_t can be deleted from the second line of (36). We take this for granted in the remainder of Appendix A.

Now, fix a period $\bar{t}\geq 0$ and a $j\leq k$, so that $N_j=\infty$. We see that in the total sum defining $\mathcal{L}_\mu$ the term $h_0^j(x_{\bar{t}},\,a_{\bar{t}},\,s_{\bar{t}})$ appears in the first line of (36) – the objective function of $\mathbf{PP}_\mu$ – premultiplied by $\beta^{\bar{t}}\mu^j$; this term also appears in the second line of (36) in the forward-looking constraints (3) at all $t\leq\bar{t}$, in the second line it is multiplied by the discounting β^n for $n=\bar{t}-t$ and then again by β^t. Therefore in the total sum in (36) $P(s^t|s_0)h_0^j(x_{\bar{t}},\,a_{\bar{t}},\,s_{\bar{t}})$ is multiplied by the following term

$$\beta^{\bar{t}}\mu^j+\gamma_0^j\beta^{\bar{t}}+\beta^1\gamma_1^j\beta^{\bar{t}-1}+...+\beta^{\bar{t}-1}\gamma_{\bar{t}-1}^j\beta^1=\beta^{\bar{t}}\left[\sum_{i=0}^{\bar{t}-1}\gamma_i^j+\mu^j\right]=\beta^{\bar{t}}\mu_{\bar{t}}^j.$$

The equalities follow from simple algebra, (6) and $j\leq k$. This gives that (4) and (5) are equivalent for $j\leq k$.

Similarly, fix $\bar{t}\geq 0$ for $j>k$ so that $N_j=0$. Then $h_0^j(x_{\bar{t}},\,a_{\bar{t}},\,s_{\bar{t}})$ for $\bar{t}>0$ appears in the first line of $\mathcal{L}_\mu$ premultiplied by $\beta^0\mu^j$ and it does not appear in the second line. For $\bar{t}>0$ the term appears once in the

and Werner (2017) follow our approach more closely and, applying their extension of the Envelope Theorem, provide a recursive formulation for the non-differentiable case.

forward-looking constraint of $\bar{t}-1$, therefore multiplied by $\beta^{\bar{t}-1}\gamma^j_{\bar{t}-1}\beta^1$. Given (6) for $j > k$ we have $\mu^j_{\bar{t}} = \gamma_{t-1}$ for $\bar{t} > 0$ and $\mu^j_0 = \mu^j$, so that the term $h^j_0(x_{\bar{t}},\, a_{\bar{t}},\, s_{\bar{t}})$ is multiplied in the total sum above by $\beta^{\bar{t}}\mu^j_{\bar{t}}$.

Hence (4) and (5) are equivalent.

Appendix B. Proofs of Theorems 1 and 2 and Proposition 1.

Proposition 1 and Theorem 1 are stated with relation to a Lagrangian that only accounts for the forward-looking constraint (3) of period zero at (x_0, s_0) – the $\mathbf{SPP}_\mu$:

$$\mathbf{SPP}_\mu: \quad SV(x,\mu,s) = \text{SP} \inf_{\gamma\in R^{l+1}_+} \sup_{\{a_t\}^\infty_{t=0}} \Big\{\mu h_0(x_0,\, a_0,\, s_0) + \gamma h_1(x_0,\, a_0,\, s_0) + \beta \; \mathrm{E}_0 \sum_{j=0}^{l} \varphi^j(\mu,\, \gamma) \sum_{t=0}^{N_j} \beta^t \; h^j_0(x_{t+1},\, a_{t+1},\, s_{t+1}) \Big\} \tag{37}$$

$$\text{s.t. } x_{t+1} = \ell(x_t,\, a_t,\, s_{t+1}), \;\; p(x_t, a_t,\, s_t) \geq 0, \quad t \geq 0, \tag{38}$$

$$\mathrm{E}_t \sum_{n=1}^{N_j+1} \beta^n h^j_0(x_{t+n},\, a_{t+n},\, s_{t+n}) + h^j_1(x_t,\, a_t,\, s_t) \geq 0,\;\; j = 0,...,l,\;\; t \geq 1, \tag{39}$$

where (37) is obtained by adding the term $\gamma \left[\mathrm{E}_0 \sum_{t=0}^{N_j+1} \beta^t \; h^j_0(x_{t+1},\, a_{t+1},\, s_{t+1}) + h_1(x_0,\, a_0,\, s_0)\right]$ to the objective function of $\mathbf{PP}_\mu$ and rearranging. Notice how (39) only holds for $t \geq 1$, that is, this Lagrangian only attaches a multiplier to the forward-looking constraint (3) at $t = 0$, while the remaining constraints (3) for $t > 0$ are kept as constraints.

Given the definition of saddle-point that we use, $(\boldsymbol{a}^*, \gamma^*_0)$ solves $\mathbf{SPP}_\mu$ at (x, s) if given $\gamma^*_0 \in R^{l+1}_+$, the path $\boldsymbol{a}^*$ is maximal for (37) respect to all the paths satisfying (38) - (39) and, given $\boldsymbol{a}^*$, γ^*_0 is a minimal element for (37) in R^{l+1}_+.

The ∞-dimensional formulation

For some of the proofs it is convenient to describe the *infinite-dimensional* formulation of $\mathbf{PP}_\mu$. The underlying uncertainty takes the form of an exogenous stochastic process $\{s_t\}^\infty_{t=0}$, $s_t \in S$, 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 $\{s_t\}^\infty_{t=0} \in S_\infty$, 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. Plans, $\boldsymbol{a} = \{a_t\}^\infty_{t=0}$, are elements of $\mathcal{A} = \{\boldsymbol{a} : \forall t \geq 0,\, a_t : S_t \to A \text{ and } a_t \in \mathcal{L}^m_\infty(S_t, \mathcal{S}_t, P)\}$, where $\mathcal{L}^m_\infty(S_t, \mathcal{S}_t, P)$ denotes the space of m-valued, essentially bounded, $\mathcal{S}_t$-measurable functions. The corresponding endogenous state variables are elements of $\mathcal{X} = \{\boldsymbol{x} : \forall t \geq 0,\, x_t \in \mathcal{L}^n_\infty(S_t, \mathcal{S}_t, P)\}$.

We now define preferences, sets of feasible actions and problems, given initial conditions (x, s). A plan $\boldsymbol{a} \in \mathcal{A}$ and a corresponding $\boldsymbol{x} \in \mathcal{X}$, is evaluated in $\mathbf{PP}_\mu$ by

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

We can describe the *forward-looking constraints* , coordinatewise, $g_{(x,s)}(\cdot)_{\,t} : \mathcal{A} \to \mathcal{L}^{l+1}_\infty(S_t, \mathcal{S}_t, P)$ by

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

The corresponding feasible set of plans is then

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

Therefore, the $\mathbf{PP}_\mu$ can be written in compact form as:

$$\mathbf{PP}_\mu \qquad \sup_{\boldsymbol{a}\in\mathcal{B}(x,s)} f_{(x,\mu.s)}(\boldsymbol{a}). \tag{40}$$

We denote solutions to this problem as $\boldsymbol{a}^*$ and the corresponding sequence of state variables as $\boldsymbol{x}^*$. When the solution exists the value function of $\mathbf{PP}_\mu$ can be written as $V_\mu(x,s)=f_{(x,\mu.s)}(\boldsymbol{a}^*)$.

It will be useful to consider a feasible set that disregards the *forward-looking constraints* in the initial period resulting in

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

Then, the $\mathbf{SPP}_\mu$ problem defined above can be written using this formulation as:

$$\mathbf{SPP}_\mu \qquad \text{SP}\inf_{\gamma\in R^l_+}\sup_{\boldsymbol{a}\in\mathcal{B}'(x,s)}\left\{f_{(x,\mu,.s)}(\boldsymbol{a})+\gamma g(\boldsymbol{a})_{0,s}\right\}.$$

Proof of Theorem 1 Part I ($\mathbf{SPP}_\mu \Rightarrow \mathbf{PP}_\mu$): It follows from Theorem 2, section 8.4 in Luenberger (1969, p. 221) that $\boldsymbol{a}^*$ solves $\mathbf{PP}_\mu$ and the value at the saddle point is the same as the value at the maximum, hence $SV(x,\mu,s)=V_\mu(x,s)$ ■

Proof of Theorem 1 Part II ($\mathbf{SPP}_\mu \Rightarrow$ SPFE): We need to show that $W(x,\mu,s)=V_\mu(x,s)=SV(x,\mu,s)$ satisfies the **SPFE** and that the period zero solution of $\mathbf{SPP}_\mu$ at $(x,\ s)$, namely (a_0^*,γ_0^*), is a saddle-point of **SPFE** at (x,μ,s).

First we show that, given γ_0^*, a_0^* satisfies the maximand part (20) for $W=SV$. 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 given initial condition $\widetilde{x}_1=\ell(x,\ \widetilde{a},\ s_1)$. 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:

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

Part I of this theorem and the definition of $\mathbf{PP}_{\mu_1^*}$ imply that

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

Since the sequence $\widetilde{\boldsymbol{a}}$ is feasible for $\mathbf{SPP}_\mu$ (that is, $\widetilde{\boldsymbol{a}}$ may fail the forward-looking constraint at $t=0$, but recall that this constraint does not constrain the feasible set in $\mathbf{SPP}_\mu$) and since $\boldsymbol{a}^*(x,\mu,s)$ solves the *sup* part of $\mathbf{SPP}_\mu$, given (41) we have the first inequality in

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

the equality follows because (41) also works when $\widetilde{\boldsymbol{a}}$ is replaced by $\boldsymbol{a}^*(x,\mu,s)$.

This proves that a_0^* satisfies (20) when $W=SV$.

To show that γ_0^* satisfies (19), note that given any $\widetilde{\gamma}\in R_+^{l+1}$,

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

where the inequality follows from the fact that the continuation of $\boldsymbol{a}^*$ is feasible but not necessarily optimal for $\mathbf{PP}_{\varphi(\mu,\,\widetilde{\gamma})}$ at (x_1^*,s_1). Using this and the fact that γ^* solves the min part of $\mathbf{SPP}_\mu$, we have

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

This proves that $(a_0^*,\gamma_0^*)\in\Psi_{SV}(x,\,\mu,\,s)$. Finally, using the definition of SV in (37) 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 Corollary to Theorem 1: We have to show that, with the additional assumptions, $(\mathbf{PP}_\mu \Rightarrow \mathbf{SPP})_\mu$ –i.e. there exists a $\gamma^*\in R_+^{l+1}$ such that $(\boldsymbol{a}^*,\gamma^*)$ is a solution to $\mathbf{SPP}_\mu$ with initial conditions (x,s). With the above formulation (40), this is an immediate application of Theorem 1 (8.3) and Corollary 1 in Luenberger (1969, p.217). To see this, note that $\mathcal{B}'(x,s)$ is a convex subset of $\mathcal{A}$, $g_{(x,s)}(\cdot)_0:\mathcal{A}\to\mathcal{L}_\infty^{l+1}(S_0,\mathcal{S}_0,P)$, and by **A7**[40]there is an $\widetilde{\boldsymbol{a}}\in\mathcal{B}'(x,s)$ such that $g_{(x,s)}(\widetilde{\boldsymbol{a}})_0>0$ ■

Proof of Proposition 1: Let $\widehat{S}_1\subset S$ be the set such that if $s_1\in\widehat{S}_1$ then

$$V_{\mu_1^*}(x_1^*,s_1)>\mathrm{E}\left[\sum_{j=0}^{l}\sum_{t=0}^{N_j}\beta^t\mu_1^{*j}h_0^j(x_{t+1}^*,\,a_{t+1}^*,\,s_{t+1})\mid s_1\right].$$

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

[40]Assumption **A7** is weaker than the standard Slater's condition but, when the concavity assumption **A6** is satisfied, it is equivalent.

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

$$\mathrm{E}\left[\sum_{j=0}^{l} \mu_1^{j*} \sum_{t=0}^{N_j} \beta^t \ h_0^j(x_{t+1}^*, \, a_{t+1}^*, \, s_{t+1}) \mid s_1 = \widehat{s}_1\right] < \mathrm{E}\left[\sum_{j=0}^{l} \mu_1^{j*} \sum_{t=0}^{N_j} \beta^t \ h_0^j(\widehat{x}_t, \, \widehat{a}_t, \, s_t) \mid s = \widehat{s}_1\right]. \quad (43)$$

Denote by $\widetilde{\boldsymbol{a}}$ an allocation such that $\widetilde{a}_0 = a_0^*$, it maintains the saddle point for $t > 0$ so $\{\widetilde{a}_t\}_{t=1}^\infty = \{a_t^*\}_{t=1}^\infty$ if $s_1 \in S \backslash \widehat{S}_1$, while the solution switches so $\{\widetilde{a}_t\}_{t=1}^\infty = \{\widehat{a}_t\}_{t=0}^\infty$ if $s_1 \in \widehat{S}_1$. If $\mathrm{Prob}(\widehat{S}_1) > 0$ we have

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

where the inequality follows from (43).

Finally, note that the plan $\widetilde{\boldsymbol{a}}$ is feasible for $\mathbf{SPP}_\mu$. This is because since $\{\widehat{a}_t\}_{t=0}^\infty$ solves $\mathbf{PP}_{\mu_1^*}$ it satisfies the constraints in (39) (note that $\widetilde{\boldsymbol{a}}$ will generically violate the forward-looking constraint at $t = 0$, but this constraint is absent in (39)). Therefore the above inequality contradicts that $\boldsymbol{a}^*$ solves the max part of $\mathbf{SPP}_\mu$ with initial conditions (x, s) and it contradicts the assumption that $(\boldsymbol{a}^*, \gamma_0^*)$ is a saddle point of $\mathbf{SPP}_\mu$. It follows that $\mathrm{Prob}(\widehat{S}_1) = 0$ or, equivalently, $V_{\mu_1^*}(x_1^*, s_1) \leq \mathrm{E}\left[\sum_{j=0}^{l} \sum_{t=0}^{N_j} \beta^t \mu_1^{*j} h_0^j(x_{t+1}^*, \, a_{t+1}^*, \, s_{t+1}) \mid s_1\right]$ a.s.

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

Therefore, $V_{\mu_1^*}(x_1^*, s_1) = \mathrm{E}\left[\sum_{j=0}^{l} \sum_{t=0}^{N_j} \beta^t \mu_1^{*j} h_0^j(x_{t+1}^*, \, a_{t+1}^*, \, s_{t+1}) \mid s_1\right]$ a.s. and $\{a_t^*\}_{t=1}^\infty$ solves $\mathbf{PP}_{\mu_1^*}$ with initial conditions (x_1^*, s_1) a.s. ■

Proof of Theorem 2 Part II: We need to show that if $(\mathbf{a}^*, \boldsymbol{\gamma}^*)_{(x,\mu,s)}$ is generated by the *saddle-point policy correspondence* Ψ_W – i.e. $(a_t^*, \, \gamma_t^*) \in \Psi_W(x_t^*, \, \mu_t^*, \, s_t)$ for every (t, s_t) – then $\mathbf{a}^*$ is a solution to $\mathbf{PP}_\mu$ at (x, s), already knowing that it satisfies the constraints of $\mathbf{PP}_\mu$. In particular, if there is a program $\{\widetilde{a}_t\}_{t=0}^\infty$, and $\{\widetilde{x}_t\}_{t=0}^\infty$, given by $\widetilde{x}_0 = x, \widetilde{x}_{t+1} = \ell(\widetilde{x}_t, \, \widetilde{a}_t, \, s_{t+1})$, satisfying the constraints of $\mathbf{PP}_\mu$ with initial condition (x, s), then this program cannot result in a higher value than $W(x, \mu, s)$. To this end, note that the maximality condition (20) can be expressed as:

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

$$\geq \mu h_0(x,\, a,\, s) + \gamma^* h_1(x,\, a,\, s) + \beta\ \mathrm{E}\left[\varphi(\mu,\, \gamma^*)\omega(x',\, \varphi(\mu,\, \gamma^*),\, s')|\, s\right]. \tag{44}$$

and

$$W(x_t^*,\, \mu_t^*,\, s_t) = \mu_t^* h_0(x_t^*,\, a_t^*,\, s_t) + \gamma_t^* h_1(x_t^*,\, a_t^*,\, s_t) + \beta\ \mathrm{E}\left[W(x_{t+1}^*,\, \varphi(\mu_t^*,\, \gamma_t^*),\, s_{t+1})|\, s_t\right]. \tag{45}$$

Furthermore, to simplify the notation, let $\mu_1^* = \varphi\left(\mu, \gamma_0^*\right), \widetilde{\mu}_2^* = \varphi(\mu_1^*, \gamma_1^*(\widetilde{x}_1))$[41] 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, (44), and we expand the value function according to (45). In particular, inequalities (46), (48) and (51) apply the inequality (44), and the equalities (47) and (49) apply the equality (45), while equality (50) simply rearranges terms and (52) uses the transversality condition, $\lim_{T\longrightarrow\infty} \beta^T W = 0$.We conclude the proof of the max part of **SPP** by showing that the left-hand side of (46) is greater or equal to (53):

$$\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] && (46) \\
=\ & \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] && (47) \\
\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] && (48) \\
=\ & \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] && (49) \\
=\ & \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] && (50) \\
& \ldots
\end{aligned}$$

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

$$\begin{aligned}
&\geq \quad A_T \equiv \mu \left[h_0(x,\ \widetilde{a}_0,\ s) + \beta I^k \mathrm{E}\left[\sum_{t=0}^{T-1} \beta^t h_0(\widetilde{x}_{t+1},\ \widetilde{a}_{t+1},\ s_{t+1}) | s \right]\right] \\
&\quad +\gamma_0^* \left[h_1(x,\ \widetilde{a}_0,\ s) + \beta \mathrm{E}\left[h_0(\widetilde{x}_1,\ \widetilde{a}_1,\ s_1) + I^k \sum_{t=1}^{T-1} \beta^t h_0(\widetilde{x}_{t+1},\ \widetilde{a}_{t+1},\ s_{t+1}) | s \right]\right] \\
&\quad +\beta \mathrm{E}\left[\gamma_1^*(\widetilde{x}_1) \left[h_1(\widetilde{x}_1,\ \widetilde{a}_1,\ s_1) + \beta \left[h_0(\widetilde{x}_2,\ \widetilde{a}_2,\ s_2) + I^k \sum_{t=2}^{T-1} \beta^{t-1} h_0(\widetilde{x}_{t+1},\ \widetilde{a}_{t+1},\ s_{t+1}) \right]\right] | s \right] \\
&\quad ... \\
&\quad +\beta^T \mathrm{E}\left[\gamma_T^*(\widetilde{x}_T) h_1(\widetilde{x}_T,\ \widetilde{a}_T,\ s_T) | s \right]
\end{aligned}$$

$$+\beta^{T+1} \mathrm{E}\left[\varphi\left(\mu_T^*, \gamma_T^*(\widetilde{x}_T)\right) \omega(\ell(\widetilde{x}_T,\ \widetilde{a}_T,\ s_{T+1}), \varphi\left(\mu_T^*, \gamma_T^*(\widetilde{x}_T)\right)\ ,\ s_{T+1}) | \, s \right], \tag{51}$$

$$\begin{aligned}
\lim_{T \longrightarrow \infty} A_T &= \mu \left[h_0(x,\ \widetilde{a}_0,\ s) + \beta I^k \mathrm{E}\left[\sum_{t=0}^{\infty} \beta^t h_0(\widetilde{x}_{t+1},\ \widetilde{a}_{t+1},\ s_{t+1}) | s \right]\right] \\
&\quad +\gamma_0^* \left[h_1(x,\ \widetilde{a}_0,\ s) + \beta \mathrm{E}\left[h_0(\widetilde{x}_1,\ \widetilde{a}_1,\ s_1) + I^k \sum_{t=1}^{T-1} \beta^t h_0(\widetilde{x}_{t+1},\ \widetilde{a}_{t+1},\ s_{t+1}) | s \right]\right] \qquad (52) \\
&= \mu h_0(x,\ \widetilde{a}_0,\ s) + \gamma_0^* h_1(x,\ \widetilde{a}_0,\ s) \\
&\quad +\beta\ \mathrm{E}\left[\sum_{j=0}^{l} \varphi^j(\mu,\ \gamma_0^*) \sum_{t=0}^{N_j} \beta^t\ h_0^j(\widetilde{x}_{t+1},\ \widetilde{a}_{t+1},\ s_{t+1}) | s \right]. \qquad (53)
\end{aligned}$$

In sum,

$$\begin{aligned}
W(x, \mu, s) \quad &\geq \quad \mu h_0(x,\ \widetilde{a}_0,\ s) + \gamma_0^* h_1(x,\ \widetilde{a}_0,\ s) \\
&\quad +\beta\ \mathrm{E}\left[\sum_{j=0}^{l} \varphi^j(\mu,\ \gamma_0^*) \sum_{t=0}^{N_j} \beta^t\ h_0^j(\widetilde{x}_{t+1},\ \widetilde{a}_{t+1},\ s_{t+1}) | s \right],
\end{aligned}$$

and, therefore, $W(x, \mu, s) = V_\mu(x, s)$ ■

Appendix C. Properties of value functions and supporting results on subdifferential calculus. Some properties of $V_\mu(x, s)$ and SPFE

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

i) $V_\mu(x,\ s)$ is convex and homogeneous of degree one in μ,

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

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

Proof: *i)* To simplify notation, denote the solution of $\mathbf{PP}_\mu$ 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

$$f_{(x, \lambda\mu + \lambda'\mu', s)}(\boldsymbol{a}) = \lambda f_{(x, \mu, s)}(\boldsymbol{a}) + \lambda' f_{(x, \mu', s)}(\boldsymbol{a}) \tag{54}$$

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

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

where the first equality follows from (54) and the inequality follows 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 (54) and the fact that $\boldsymbol{a}_{\lambda\mu}^{*}$ and $\boldsymbol{a}_{\mu}^{*}$ are maximal elements attaining $V_{\lambda\mu}(x,\, s)$ and $V_{\mu}(x,\, s)$ respectively:

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

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, its value 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, support our discussion on 'uniqueness and sufficiency without differentaibility' in Section 3 and, in particular, are used in proving Lemma 1 and Lemma 5A (below):

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 $F^i : R^m_+ \to R$ 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 $F(x) - F(x^*) \geq 0$ $\blacksquare$

Sufficiency (without differentiability): supporting results

Lemma 5A. 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, μ, 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:

$$\begin{aligned}&\mu h_0(x, a^*, s) + \gamma^* h_1(x, a^*, s) + \beta\, \mathrm{E}\left[\varphi(\mu, \gamma^*)\omega(x^{*\prime}, \varphi(\mu, \gamma^*), s') | \, s\right] \\ &\geq \mu h_0(x, a, s) + \gamma^* h_1(x, a, s) + \beta\, \mathrm{E}\left[\varphi(\mu, \gamma^*)\omega(x', \varphi(\mu, \gamma^*), s') | \, s\right],\end{aligned} \tag{55}$$

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

$$h_1^j(x, a^*, s) + \beta\, \mathrm{E}\left[\omega^j(x^{*\prime}, \varphi(\mu, \gamma^*), s') | \, s\right] \geq 0, \tag{56}$$

$$\gamma^{*j}\left[h_1^j(x, a^*, s) + \beta\, \mathrm{E}\left[\omega^j(x^{*\prime}, \varphi(\mu, \gamma^*), s') | \, s\right]\right] = 0. \tag{57}$$

Proof: Part (***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 (***ii***) (55) is the same as the max saddle-point condition of **SPFE** expressed with its Euler representation. Since by (***i***) W always has at least one Euler representation, the proof of (55) is immediate. To see the min inequality of part (***ii***), begin by rewriting the first inequality of $\Psi_W(x,\, \mu,\, s)$, (19, 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 (56) and (57) are necessary and sufficient ■

Appendix D. Proof of Theorem 3.

The proof of Theorem 3(i) is based on the following two lemmas and Kakutani's fixed point theorem. .

Lemma 6A. Assume **A4** and that $W \in \mathcal{M}_{bc}$ satisfies **SIC**. There exists a $C > 0$ such that if $(a^*, \gamma^*) \in \Psi_W(x, \mu, s)$ then $\|\gamma^*\| \leq C\, \|\mu\|$.

Before we prove Lemma 6A, note that condition **SIC** implies the following condition, which is a version of *Karlin's condition*[42] :

SK. W, with $W = \mu\omega$, satisfies the *interiority condition* if there exists an $\epsilon > 0$, such that for any $(x,\, s) \in X \times S$, $\mu \in R_+^{l+1}$, and $\gamma \in R_+^{l+1}, \gamma \neq 0$ there exists $\widetilde{a} \in A$, satisfying $p(x,\, \widetilde{a},\, s) > 0$, and $\gamma\left[h_1(x, \widetilde{a}, s) + \beta \mathrm{E}\left[\omega(\ell(x,\, \widetilde{a},\, s'),\, \mu, s')|\, s\right]\right] \geq \epsilon$.

Proof: Lemma 6A is trivially satisfied if $\gamma^* = 0$, therefore let $(a^*, \gamma^*) \in \Psi_W(x, \mu, s)$ with $\gamma^* \neq 0$ and the Euler representations $W(\ell(x,\, a^*,\, s'),\, \varphi(\mu,\, \gamma^*),\, s') = \mu'^*\omega(x'^*, \mu'^*, s')$, and let $\widetilde{a} \in B(x, s)$ be the interior allocation of the **SIC** condition. Using the same notation as in the proof of Theorem 4, the slackness condition $\gamma^*\left[h_1(x,\, a^*,\, s) + \beta\ \mathrm{E}\left[\omega(x'^*,\, \mu'^*,\, s')|\, s\right]\right] = 0$ and **SIC,** the max inequality can be expressed as:

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

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

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

[42] See Takayama (1985).

which, by **SIC**, 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 $\|\gamma^*\| \leq C\|\mu\|$ ■

The next lemma requires some additional notation. Let $B(x,s)=\{a\in A: \; p(x,a,s)\geq 0\}$, and $G(\mu)=\left\{\gamma\in R_+^{l+1}: \|\gamma\|\leq C\|\mu\|\right\}$, where $\|\mu\|>0$. Define

$$SP^a_{W(x,\,\mu,\,s)}(\gamma)=\left\{\begin{array}{c} a\in B(x,s): \mu h_0(x,\,a,\,s)+\gamma h_1(x,\,a,\,s)+\beta \mathrm{E}\left[W(\ell(x,\,a,\,s'),\,\varphi(\mu,\,\gamma),\,s')|\,s\right] \\ \geq \mu h_0(x,\,\widetilde{a},\,s)+\gamma h_1(x,\,\widetilde{a},\,s)+\beta \mathrm{E}\left[W(\ell(x,\,\widetilde{a},\,s'),\,\varphi(\mu,\,\gamma),\,s')|\,s\right],\,\forall\,\widetilde{a}\in B(x,s)\end{array}\right\},$$

$$SP^\gamma_{W(x,\,\mu,\,s)}(a)=\left\{\begin{array}{c} \gamma\in G(\mu): \mu h_0(x,\,a,\,s)+\gamma h_1(x,\,a,\,s)+\beta \mathrm{E}\left[W(\ell(x,\,a,\,s'),\,\varphi(\mu,\gamma),\,s')|\,s\right] \\ \leq \mu h_0(x,\,a,\,s)+\widetilde{\gamma} h_1(x,\,a,\,s)+\beta \mathrm{E}\left[W(\ell(x,\,a,\,s'),\,\varphi(\mu,\,\widetilde{\gamma}),\,s')|\,s\right],\,\forall\,\widetilde{\gamma}\in G(\mu)\end{array}\right\},$$

and $SP_{W(x,\,\mu,\,s)}: B(x,s){\times}G(\mu)\to B(x,s){\times}G(\mu)$ by $SP_{W(x,\,\mu,\,s)}(a,\gamma)=\left(SP^a_{W(x,\,\mu,\,s)}(\gamma), SP^\gamma_{W(x,\,\mu,\,s)}(a)\right)$.

Lemma 7A. Assume **A1-A5** and that $W\in\mathcal{M}_{bc}$ satisfies **SIC.** The correspondence $SP_{W(x,\,\mu,\,s)}$ is nonempty, convex-valued and upper hemi-continuous.

Proof: $SP_{W(x,\,\mu,\,s)}$ is a max and min problem of continuous functions on compact sets with non-empty interiors and, therefore, for all $(a,\gamma)\in B(x,s)\times G(\mu)$ is non-empty and, given our concavity assumptions it is convex-valued. To see that it is upper hemi-continuous let $(a_n,\gamma_n)\to(a,\gamma)$ with $a_n\in SP^a_{W(x,\,\mu,\,s)}(\gamma_n)$ and $\gamma_n\in SP^\gamma_{W(x,\,\mu,\,s)}(a_n)$ – i.e. for all $\widetilde{a}\in B(x,s)$ $a_n\succsim\widetilde{a}$ and for all $\widetilde{\gamma}\in G(\mu)$ $\gamma_n\precsim\widetilde{\gamma}$,for all n, but, by continuity of the implied functions, $a\succsim\widetilde{a}$ and $\gamma\precsim\widetilde{\gamma}$, therefore $(a,\gamma)\in SP_{W(x,\,\mu,\,s)}(a,\gamma)$ ■

Proof of Theorem 3(i): The assumptions of Lemmas 6A and 7A are also assumed in Theorem 3(i), therefore the correspondence $SP_{W(x,\,\mu,\,s)}: B(x,s)\times G(\mu)\to B(x,s)\times G(\mu)$ mapping nonempty, convex and compact sets to themselves, is nonempty, convex-valued and upper-hemi continuous and by Kakutani's Fixed Point Theorem (e.g. Mas-Colell *et al.* (1995)) there exists $(a^*,\gamma^*)\in SP_{W(x,\,\mu,\,s)}(a^*,\gamma^*)$. Finally, that a^* is unique when $W\in\mathcal{M}_{bc}$ and **A6s** is satisfied, is a standard result (see footnote (23) ■

Before we prove 3(ii) note that, given the assumptions of Theorem 3, $(a^*,\gamma^*)\in SP_{W(x,\,\mu,\,s)}(a^*,\gamma^*)$ if and only if $(a^*,\gamma^*)\in\Psi_W(x,\mu,s)$. If $(a^*,\gamma^*)\in SP_{W(x,\,\mu,\,s)}(a^*,\gamma^*)$ then (a^*,γ^*) satisfies inequalities (19) - (20), with the former restricted to $G(\mu)$, but by Lemma 6A this restriction is not binding once **SIC** is assumed. Conversely, if $(a^*,\gamma^*)\in\Psi_W(x,\mu,s)$ then $a^*\in SP^a_{W(x,\,\mu,\,s)}(\gamma^*)$ and $\gamma^*\in SP^\gamma_{W(x,\,\mu,\,s)}(a^*)$. Obviously, only when *saddle-point solutions* are unique – i.e. $(a^*,\gamma^*)=\psi_W(x,\mu,s)$ – we have $\psi_W(x,\mu,s)$ $=SP_{W(x,\mu,s)}(a^*,\gamma^*)=(a^*,\gamma^*)$.

Proof of Theorem 3(ii): First we show that, given $W\in\mathcal{M}$, $T^*W(\cdot,\cdot,s)$ is also continuous by extending the *Theorem of the Maximum* to *saddle-points.* That $SP_{W(x,\,\mu,\,s)}$ satisfies the closed-graph property (Lemma 7A) implies that $\Psi_W(x,\mu,s)\subset B(x,s)\times G(\mu)$ is closed. Furthermore, $B(\cdot,s): X\longrightarrow A$ and $G(\cdot): R_+^{l+1}\to R_+^{l+1}$ are continuous correspondences. Now to show that Ψ_W is an upper-hemi continuous correspondence, fix (x,μ) and let the sequence $(x_n,\mu_n)\to(x,\mu)$ and $(a^*_n\;,\gamma^*_n)\in\Psi_W(x_n,\mu_n,s)$, for all n. Since $B(\cdot,s)$ and $G(\cdot)$ are upper-hemi continuous, there exists a subsequence $(a^*_{n_k}\;,\gamma^*_{n_k})\to(a^*\;,\gamma^*)\in B(x,s)\times G(\mu)$ with $(a^*_{n_k}\;,\gamma^*_{n_k})\in\in\Psi_W(x_{n_k},\mu_n,s)$. Given an arbitrary $(\widehat{a},\widehat{\gamma})\in B(x,s)\times G(\mu)$, since $B(\cdot, s)$ and $G(\cdot)$ are lower-hemi continuous, there exists a subsequence $(\widehat{a}_{n_k}, \widehat{\gamma}_{n_k}) \to (\widehat{a}, \widehat{\gamma})$ with $(\widehat{a}_{n_k}, \widehat{\gamma}_{n_k}) \in B(x_{n_k}, s) \times G(\mu_{n_k})$; that is:

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

and by continuity

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

therefore $(a^*, \gamma^*) \in \Psi_W(x, \mu, s)$. Now we can show that $T^*W(\cdot, \cdot, s)$ is continuous. Again let the sequence $(x_n, \mu_n) \to (x, \mu)$ and $(a^*_n, \gamma^*_n) \in \Psi_W(x_n, \mu_n, s)$, for all n, then

$$\begin{aligned} T^*W(x_n, \mu_n, s) &= \mu_n \left[h_0(x_n,\, a^*_n,\, s) + \beta\ \mathrm{E}\left[I^k \omega(\ell(x_n,\, a^*_n,\, s'),\ \varphi(\mu_n,\, \gamma^*_n),\, s') |\, s\right]\right] \\ &\quad + \gamma^*_n \left[h_1(x_n,\, a^*_n,\, s) + \beta\ \mathrm{E}\left[\omega(\ell(x_n,\, a^*_n,\, s'),\ \varphi(\mu_n,\, \gamma^*_n),\, s') |\, s\right]\right] \\ &= \mu_n \left[h_0(x_n,\, a^*_n,\, s) + \beta\ \mathrm{E}\left[I^k \omega(\ell(x_n,\, a^*_n,\, s'),\ \varphi(\mu_n,\, \gamma^*_n),\, s') |\, s\right]\right]. \end{aligned}$$

Since the last equality is satisfied for every sequence and subsequence we only need to consider the last term. In particular, if we let $\overline{W} = \limsup T^*W(x_n, \mu_n, s)$ and $\underline{W} = \liminf T^*W(x_n, \mu_n, s)$, then there is a subsequence $\{x_{n_k}, \mu_{n_k}\}$ such that

$$\overline{W} = \lim \mu_{n_k} \left[h_0(x_{n_k},\, a^*_{n_k},\, s) + \beta\ \mathrm{E}\left[I^k \omega(\ell(x_{n_k},\, a^*_{n_k},\, s'),\ \varphi(\mu_{n_k},\, \gamma^*_{n_k}),\, s') |\, s\right]\right]$$

and, by the upper-hemi continuity of Ψ_W there is a further subsequence $(a^*_{n_{k_r}}, \gamma^*_{n_{k_r}}) \in \Psi_W(x_{n_{k_r}}, \mu_{n_{k_r}}, s)$ converging to $(a^*, \gamma^*) \in \Psi_W(x, \mu, s)$, therefore $\overline{W} = \lim T^*W\left(x_{n_{k_r}}, \mu_{n_{k_r}}, s\right) = T^*W(x, \mu, s)$. Since the same argument applies to $\underline{W}$ it follows that $T^*W(x_n, \mu_n, s) \to T^*W(x, \mu, s)$. We now show that the remaining properties of $\mathcal{M}$ are preserved.

That T^*W is also bounded follows from **A3** - **A4** and the boundedness condition on W. Furthermore, by **A1b** T^*W is measurable; therefore, it satisfies (i) of the definition of $\mathcal{M}_b$. To see that T^*W is homogeneous of degree one – i.e.$T^*W(x,\, \lambda\mu,\, s) = \lambda T^*W(x,\, \mu,\, s)$, for any $\lambda > 0$ – let $(a^*,\, \gamma^*)$ be a solution to the *saddle-point* Bellman equation at $(x,\, \mu,\, s)$ – i.e. $(a^*,\, \gamma^*) \in \Psi_W(x,\, \mu,\, s)$. It is enough to show that, for any $\lambda > 0, (a^*,\, \lambda\gamma^*) \in \Psi_W(x,\, \lambda\mu,\, s)$ – i.e. $\gamma^*(x, \lambda\mu, s) = \lambda\gamma^*(x,\, \mu,\, s)$ – since

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

and $W(x^{*\prime},\, \varphi(\lambda\mu,\, \lambda\gamma^*),\, s') = \lambda W(x^{*\prime},\, \varphi(\mu,\, \gamma^*),\, s')$. 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

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

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

Proof of Theorem 3(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 8A - 10A show that T^* satisfies the conditions of the *Contraction Mapping Theorem* and *Blackwell's sufficiency conditions* ■

Lemma 8A. $\mathcal{M}$ 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 $(x,\, \mu,\, s)$ and $\lambda > 0$,

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

Lemma 9A (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,

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

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 10A (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

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

REFERENCES

Ábrahám, A., E. Carceles-Poveda, (2010): "Endogenous trading constraints with incomplete asset markets" *Journal of Economic Theory,* 145(3), 974.

Ábrahám, A., E. Carceles-Poveda, Y. Liu and R. Marimon (2018): "On the optimal design of a Financial Stability Fund," ADEMU WP 2018/105.

Ábrahám, A. and S. Laczó (2018): "Efficient risk sharing with limited commitment and storage" *Review of Economic Studies*, vol. 85(3), 1389-1424.

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.

Adam, K. and R.M. Billi (2006): "Optimal Monetary Policy under Commitment with a Zero Bound on Nominal Interest Rates", *Journal of Money, Credit and Banking* 38(7), 1877-1905.

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

Alvarez, F. and N. Stokey (1998): "Dynamic Programming with Homogeneous Functions", *Journal of Economic Theory* 82, 167-189.

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.

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.

judd, K.L, S. yeltekin and J. Conklin (2006): "Computing Supergame Equilibria", *Econometrica* 71, 1239-1254.

Faraglia E., A. Marcet, R. Oikonomou and A. Scott (2019): "Government Debt Management: the Long and the Short of It", *Review of Economic Studies*.

Faraglia E., A. Marcet, R. Oikonomou and A. Scott (2016): "Long Term Government Bonds" working paper, BGSE.

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 (2018): *Recursive Macroeconomic Theory* (Fourth 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. 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 J. Werner (2017): "The Envelope Theorem, Euler and Bellman Equations, without Differentiability," mimeo, Univeristy of Minnesota and EUI.

Mas-Colell, A., M.D. Whinston and J.R. Green (1995): *Microeconomic Theory.* Oxford: Oxford University Press.

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.

Schmitt-grohé and M. Uribe (2004): "Optimal fiscal and monetary policy with sticky prices" *Journal of Economic Theory,* 114 (2), 198-230

Siu, H.E., (2004): "Optimal fiscal and monetary policy with sticky prices", *Journal of Monetary Economics*, 51 (3) 575-607.

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. (1985): *Mathematical Economics, 2nd edition.* Cambridge University Press.

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