RECURSIVE CONTRACTS

Albert Marcet
Universitat Pompeu Fabra, Barcelona
and
Ramon Marimon
European University Institute, Florence
and
Universitat Pompeu Fabra, Barcelona

First version: September, 1994
This version: October, 1998*

*We would like to thank Fernando Alvarez, Jean-Michel Grandmont, Edward Green, Esther Hauk, Robert Lucas, Andreu Mas-Colell, Edward Prescott, Victor Rios,Thomas Sargent, Robert Townsend for comments on earlier developments of this work.

Authors' Affiliations

Albert Marcet
Universitat Pompeu Fabra, Departament d'Economia i Empresa,
Ramon Trias Fargas 25-27, 08005 Barcelona, Catalonia, Spain;
Telephone: +34-93-542-2740, Fax: +34-93-542-1746,
E-mail: albert.marcet@econ.upf.es
http://www.econ.upf.es/~marcet

and

Ramon Marimon
European University Institute, Economics Department
Badia Fiesolana, via dei Roccettini 9
50016 San Domenico di Fiesole (FI), Italy,
Telephone: +39-055-468-5252, Fax: +39-055-468-5202
E-mail: marimon@datacomm.iue.it
http://www.iue.it/Personal/Marimon/Welcome.html.

Keywords: Time inconsistency, recursive formulation, dynamic programming, Ramsey equilibrium, participation constraint, recursive saddle points.
JEL classification: C61, C63.

Abstract

We obtain a recursive formulation for a general class of contracting problems involving incentive constraints. Under these constraints, the corresponding maximization (*sup*) problems fails to have a recursive solution. Our approach consists of studying the Lagrangian. We show that, under standard assumptions, the solution to the Lagrangian is characterized by a recursive *saddle point* (*infsup*) functional equation, analogous to Bellman's equation. Our approach applies to a large class of contractual problems. As examples, we study the optimal policy in a model with intertemporal participation constraints (which arise in models of default) and intertemporal competitive constraints (which arise in Ramsey equilibria).

1 Introduction

The use of recursive analysis is one of the main resources available today to economists studying dynamic models. In the standard case, it is well known how to determine if a model has a recursive structure; for example, Stokey, *et al.* (1989) describe a large number of models that can be analyzed recursively. The presence of a recursive formulation implies that the optimal decision at time t is a time-independent function f of a small set of state variables. This property plays a crucial role in many applications of dynamic models for several reasons: it facilitates the analysis and empirical testing of the model; it is enough to approximate just one function in order to compute the equilibria for all periods[1]; contracts can be specified without taking into account all past and present realizations of exogenous stochastic shocks (as they would with Arrow-Debreu contracts) since a few state variables are sufficient statistic for past history; finally, models of learning can be formulated by specifying f as the object to be learned.

A key condition in standard dynamic programming techniques is that only past variables can influence the set of feasible current actions. Kydland and Prescott (1977) showed that many dynamic economic models of interest, failed to satisfy this condition and, therefore, the Bellman equation failed to hold in these models. This is a well known problem in dynamic games where it is usually imposed that an equilibrium solution must be *sub-game perfect*. This lack of recursivity is likely to arise in contracting problems, where intertemporal participation, incentive or competitive constraints define the set of feasible contracts. Also, in models of optimal policy, agent's reactions to government policies are taken as constraints. In all those cases, future actions limit the set of current feasible actions available to the planner. Despite the increased interest in the study of optimal dynamic contracting problems, a general method for finding a proper recursive formulation is still absent from the literature.

In this paper we provide an integrated approach for a recursive formulation of a large class of economic models. We show how, in many cases where implementability constraints depend on plans for future variables and the original maximization problem is not recursive, an equivalent recursive

[1] Several computational algorithms that exploit the recursive structure of the solution are decribed, for example, in the volumes of Cooley (1995) and Marimon and Scott (1998).

saddle point problem can be constructed leading to a recursive formulation.

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

The resulting saddle point problem (SPP) expands the set of state variables to include new variables that summarize the evolution of the lagrange multipliers of the original (PP) problem[2]. Such transformation creates some technical difficulties since the new (co)state variables can not be bounded. Fortunately, we can exploit the resulting homogeneity properties of the return function and, in this way, we are able to extend the standard contraction mapping approach to establish the relationship between SPP and the SPFE.

We show that solving the lagrangean (SPP) is equivalent to solving the recursive SPFE without concavity assumptions. This is important because incentive constraints may not have a convex structure. If concavity is satisfied, then solving the SPP (and, therefore, the SPFE) is equivalent with solving the maximization problem PP. In the absence of concavity, as in any application of lagrangean theory, our SPFE characterization is sufficient but it may not be necessary for a solution.[3]

[2]With this formulation, the resulting stationary policy function is continuous. Without the additional (co)state variables, the value function would be discontinuous (to account for the fact that non incentive compatible paths are unfeasible). Rustichini (1996) has recently followed the approach of allowing discontinuities of the value function. He does not add co-state variables. Unfortunately, this approach is very limited since, as our work shows, new co-state variables that account for when and how incentive constraints have been binding need to be introduced in order to achieve the optimum under full commitment.

[3]Most of the literature on Ramsey taxation proceeds to analyze solutions to the lagrangean even in the absence of concavity. The lagrangean approach does work in practice in most cases, since it happens that the lagragean often has a solution. To the extent that our recursive formulation characterizes all solutions to the SPP, concavity is no more necessary for our approach than it is for standard applications of Ramsey equilibria.

Our approach can be applied to a very large class of dynamic macroeconomic models, such as models of optimal fiscal or monetary policy, business cycle or financial markets models with default. It can also be used in industrial organization models of optimal regulation, game theoretical models as well as the study of enforceability and monitoring of contracts subject to incentive constraints. We do maintain, in this paper, the assumption of full information[4].

That a large class of problems have a common recursive structure is not just a technical result. It also helps in providing a common economic characterization of many contractual problems, perhaps, similar to the way that the study of recursive competitive equilibria has enhanced our understanding of the economic structure that is common to many dynamic economic models. For example, a standard application of the Second Welfare Theorem shows that, under some assumptions, recursive competitive equilibria are solutions to a planner's problem in which agents' weights are constant across time. Our recursive characterization makes it clear that, with intertemporal incentive constraints, the recursive solutions correspond to planner's problems where agents' weights vary according to their histories (more precisely, according to how incentive constraints have been binding in the past). In our approach, the additional (co)state variables indicate whether and how such adaptation of the planner's objective function must take place. In general, such adaptation of the planner's objective not only can affect the relative weight across agents, but also the planner's intertemporal valuations. For example, time-inconsistency problems can be interpreted as the planner's temptation to set the (co)state variables to its initial zero value.

The fact that at the initial period our (co)state variables are well defined (in fact, they are zero, reflecting the fact that there is no past history) allows for a proper recursive formulation. This, for example, is not true of the approach of taking "present values as state variables," pioneered by the work on repeated games of Abreu, *et al.* (1990), since initial present values of an optimal problem can only be obtained once the problem has been solved. In practice, it is often difficult to bound the range of possible initial values, as a backwards iteration of "future present values" of the APS approach requires. Furthermore, in order to find the optimal allocation under the im-

[4]In a follow up paper we characterize recursive contracts with incentive constraints under private information.

plementability constraints we do not need to characterize the whole set of feasible contracts (of all *sub-game perfect equilibria* in repeated games), and we can exploit efficiency properties in order to obtain our recursive characterizations. The fact that we have a properly defined initial condition and that we can calculate the optimum "directly" highly simplifies the practical application of our approach.

The rest of the paper is organized as follows. In Section 2 we summarize our approach and we show how it can be applied to two simple examples[5]. Section 3 develops the *saddle-point* theory, Sections 4 its dynamic programming formulation, and Section 5 concludes. Most proofs are contained in the final appendices.

2 Formulating contracts as recursive Saddle Point Problems

In this section we summarize our approach. We discuss the relationship between the maximization problem of interest with a saddle point version of the Bellman equation and conclude that, if the set of state variables is expanded to include some new co-state variables, the problem becomes recursive. We also discuss the relationship between these state variables and the evolution of the distribution of wealth, as well as the relationship to the time-consistency problem. We apply the results to two examples. All the proofs, formalities, and the technical assumptions needed are discussed in Sections 3 and 4.

The standard case of dynamic programming is concerned with problems that take the following form (see, for example, Stokey, *et al.* (1989) and Cooley, (1995)):

Program 0

$$\begin{aligned} \sup_{\{a_t\}} \quad & \mathrm{E}_0 \sum_{t=0}^{\infty} \beta^t r(x_t,\, a_t,\, s_t) \\ \text{s.t.} \quad & x_{t+1} = \ell(x_t,\, a_t,\, s_{t+1}), \quad a_t \in A(x_t,\, s_t),\ t \geq 0 \qquad \text{(1a)} \end{aligned}$$

[5]Section 2 is practically self-contained. This should allow the potential user to apply our approach without having to go through the technicalities in the rest of the paper.

$$x_0,\ s_0 \text{ given}$$
$$a_t \text{ measurable with respect to } (\ldots, s_{t-1}, s_t),$$

where r is a given return function; $\beta \in (0, 1)$ the discount factor; $\{s_t\}$ an exogenous Markov stochastic process; x an endogenous state variable; a a control or decision variable, subject to the technological constraint A, and, finally, the transition function ℓ defines the evolution of the endogenous state. The objects β, r, A, ℓ and the transition of s_t are assumed to be known.

Under standard assumptions, this problem is known to have a recursive structure, in the sense that there exists a value function v satisfying the Bellman functional equation

$$v(x, s) = \sup_{a \in A(x,s)} \quad \{r(x, a, s) + \beta \mathrm{E}\left[v(x',\, s')|s\right]\}$$
$$\text{s.t. } x' = \ell(x,\ a,\ s')$$

This functional equation can be derived using standard dynamic programming techniques (see, for example, Stokey, *et al.* (1989)). It yields a stationary policy function f such that the optimal allocation satisfies $a_t = f(x_t, s_t)$ for all t. The key aspects of this observation are that the policy function f is the same in all periods, and that only the values of (x_t, s_t) matter from the whole past history. Given this recursive structure solving the model amounts to solving for the function f. A number of computational techniques are available for this purpose.

Nevertheless, many interesting economic problems are *not* of the form of *Program 0.* This often happens in maximization problems that include different incentive or intertemporal constraints that can not be reduced to the above technological constraint (1a). In this paper we will consider problems that can be represented in the following form:

Program 1

$$\sup_{\{a_t\}} \quad \mathrm{E}_0 \sum_{t=0}^{\infty} \beta^t r(x_t,\ a_t,\ s_t)$$
$$\text{s.t.} \quad x_{t+1} = \ell(x_t,\ a_t,\ s_{t+1}),\ \ a_t \in A(x_t,\ s_t), \qquad (1a)$$
$$g_1^j(x_t,\ a_t,\ s_t) + \mathrm{E}_t \sum_{n=1}^{N_j} \beta^n g_2^j(x_{t+n},\ a_{t+n},\ s_{t+n}) \geq 0,\ j = 1, ..., k;\ t \geq 0 \qquad (2)$$

x_0, s_0 given.

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

Clearly, we have just added the constraint (2) for given g mappings. Constraints of the form (2) are not a special case of (1a), since they involve expected values of future variables[6]. We know from Kydland and Prescott (1977) that, under these constraints, the usual Bellman equation is not satisfied, the solution is *not* of the form $a_t = f(x_t, s_t)$ for all t and the whole history of past shocks s_t can matter for today's optimal decision.

In this paper we show that problems in the canonical form of *Program 1* can also be cast in an alternative recursive framework. We consider the two canonical cases $N_j = \infty$ and $N_j = 1$; other cases can be easily incorporated. The first step in our approach is to convert *Program 1* into a recursive *saddle point problem* (SPP) of the form:

Program 2

$$\begin{aligned}
\inf_{\{\gamma_t\}} \sup_{\{a_t\}} \quad & \mathrm{E}_0 \sum_{t=0}^{\infty} \beta^t h(x_t,\, a_t,\, \mu_t,\, \gamma_t,\, s_t) \\
\text{s.t.} \quad & x_{t+1} = \ell(x_t,\, a_t,\, s_{t+1}), \quad a_t \in A(x_t,\, s_t), && \text{(1a)} \\
& \mu_{t+1} = \varphi(\mu_t,\, \gamma_t,\, s_{t+1}), \quad \gamma_t \geq 0,\ t \geq 0 && \text{(3)} \\
& \mu_0 = 0, \qquad x_0, s_0,\ \text{given}, \\
& (a_t, \gamma_t)\ \text{measurable with respect to}\ (\ldots,\, s_{t-1},\, s_t),
\end{aligned}$$

where the mappings defining the technological constraints (1a) are as before, and the mappings h, φ can be derived from r, g, N_j. Here, μ_t acts as a co-state variable, and we will show that its transition function φ depends on whether the j constraint (2) has $N_j = 1$ or ∞. Notice that SPP shares with *Program 0* the features of not having future variables in the constraints and that all the functions in the constraints are known. This is why, anticipating results, we call this a *recursive* saddle point problem. The problem where μ_0 is arbitrary is of theoretical interest, and it will be considered in Section 3 when we derive a recursive formulation.

[6]When $g^2 = r$ and $N = \infty$ this problem can not be alleviated by substituting the discounted sum of (2) with the value function $v(x_t, s_t)$. Even then, this constraint is not a special case of (1a) because the requirement that ℓ and A be known mappings in *Program 0* would be violated.

We will treat *Program 2* as a primitive program and we will show that there is a *duality* theorem linking it with *Program 1*. Unfortunately, since *Program 2* is a saddle point problem, the standard theory of dynamic programming does not apply. A main contribution of this paper is to extend dynamic programming theory to recursive saddle point problems. We show that, under certain assumptions, solutions to *Program 2* obey a *saddle point functional equation* **(SPFE)** in the sense that there exists a unique value function $W(x,\ \mu,s)$ satisfying

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

for all $(x,\,\mu,\,s)$ and such that $W(x_0,\,\mu_0,\,s_0)$ is the value of *Program 2* for initial conditions $(x_0,\,\mu_0,\,s_0)$. This is a generalization of Bellman's equation. Letting ψ be the policy correspondence of this **SPFE**, in the sense that

$$\begin{aligned} \psi(x,\,\mu,\,s) \;\; &\in \;\; \arg\inf_{\gamma\geq 0}\sup_{a\in A(x,s)} \left\{h(x,\,a,\,\mu,\,\gamma,\,s)+\beta\ \mathrm{E}\left[w(x',\,\mu',\,s')|\,s\right]\right\}, \\ & \quad\ \text{s.t. } x'=\ell(x,\,a,\,s) \\ & \quad\ \text{and } \mu'=\varphi(\mu,\,\gamma,\,s') \end{aligned}$$

The key result in this paper is that the optimal solution of *Program 1* satisfies $(a_t,\gamma_t)\ =\ \psi(x_t,\mu_t,s_t)$ for all t and $\mu_0=0$. and it implies that the solution is recursive in the sense that only the values of (x_t,μ_t,s_t) are relevant from past history and the policy function ψ is time invariant and can be found by studying the SPFE.

Notice that, in order to have the solution of the SPP equivalent with the solution of the problem of interest *Program 1* calls for setting $\mu_0=0$, while in future periods μ_t is determined according to ψ and φ. This is a special feature of the optimal plan that provides a clear interpretation of the time-inconsistency problem. It is technologically feasible to the planner to reset $\mu_t=0$ at any time t, and this is what it would do if it could ignore past commitments. But if the planner sets $\mu_t=0$ it will achieve a suboptimal allocation. Full commitment on the part of the planner means, precisely, that it commits to the evolution of μ determined by ψ and φ for all periods.

Dependence of the optimal solution on μ is the reason that the model is not recursive in the standard sense of having a time-invariant policy function of (x, s).

2.1 Present value constraints.

We now discuss how *Program 1* can be transformed into a *Program 2* formulation for the case of $N = \infty$. This means that future variables enter in the implementability constraint (2) in the form of a discounted present value.

A case of particular interest is an economy with J agents, each agent having instantaneous utility function u_j, and *intertemporal participation constraints* of the form

$$\mathrm{E}_t \sum_{n=0}^{\infty} \beta^n \; u_j(c_{j,t+n}) \geq \phi_j(\omega_t) \qquad \text{for all } j, t \tag{4}$$

These constraints restrict the utility of all agents to be at least as large as some default value $\phi_j(\omega_t)$. The planner's problem that allocates resources efficiently subject to individual participation constraints is of the form of *Program 1* with r representing the one-period social welfare function $\sum_J \alpha_j u_j$; $(\alpha_j \geq 0, \sum_J \alpha_j = 1)$, with $g_{2,j} \equiv u_j(c_{j,t})$ and $g_{1,j} \equiv u_j(c_{j,t}) - \phi_j(\omega_t)$. The solution to this problem defines a *social contract* that takes into account, not only technological, but also incentive and legal constraints[7].

Now we obtain the corresponding SPP of the form of *Program 2* for *Program 1.* Notice that the corresponding Lagrangian with respect to (2) is,

$$L \equiv \mathrm{E}_0 \sum_{t=0}^{\infty} \beta^t \left[r(x_t, a_t, s_t) + \gamma_t \left(g_1(x_t,\, a_t,\, s_t) + \mathrm{E}_t \sum_{n=0}^{\infty} \beta^n g_2(x_{t+n}, a_{t+n}, s_{t+n}) \right) \right]$$

subject to (1a), measurability constraints, and given $\gamma_t \geq 0$, where, $\beta^{-t}\gamma_t$ is the Lagrange multiplier of (2) at t.

This is still not of the form of *Program 2* above, since future variables are present in the return function of L. However, under the measurability restriction, the law of iterated expectations implies that the conditional expectations E_t in the objective function of L can be imbedded in E_0. Finally,

[7] Restrictions on budget constraints can also be written as a special case of constraints (2) for $N = \infty$. See, for example, Marcet, *et al.* (1996).

reordering terms using simple algebra, it is easy to find that the function H defined as

$$H \equiv \mathrm{E}_0 \sum_{t=0}^{\infty} \beta^t \left[r(x_t,\, a_t\, s_t) + \gamma_t g_1(x_t,\, a_t,\, s_t) + \mu_t g_2(x_t,\, a_t,\, s_t) \right]$$

$$\mu_0 \;= 0,\;\text{ and, for all } t \geq 0,\; \mu_{t+1} \;=\; \mu_t + \gamma_t$$

is such that, for all feasible sequences, $L \equiv H$..

Clearly, the saddle point of H is a special case of *Program 2*, taking

$$\begin{aligned} & h(x,\, a, \mu,\, \gamma,\, s) \\ \equiv\;\; & h_0(x,\, a,\, s) + \gamma h_1(x,\, a,\, s) + \mu h_2(x,\, a,\, s) \\ \equiv\;\; & r(x,\, a,\, s) + \gamma g_1(x,\, a,\, s) + \mu g_2(x,\, a,\, s) \end{aligned}$$

$$\varphi(\mu,\, \gamma,\, s) \equiv \mu + \gamma, \qquad\qquad \mu_0 = 0$$

Example 1. A partnership with limited commitment

We consider, as an example, a model of a *partnership*, where several agents can share their individual risks and jointly invest in a project which can not be undertaken by single (or subgroups of) agents. Formally, there is a single good and J infinitely-lived consumers, with preferences represented by $E_0 \sum_{t=0}^{\infty} \beta^t\; u(c_{j,t})$; u strictly concave and monotone; c represents individual consumption. Agent j receives an endowment of consumption good $\omega_{j,t}$ at time t. Total production is given by $F(k, \theta)$, and it can be split into consumption c and investment i. The stock of capital k depreciates at the rate δ. The joint process $\{\theta_t, \omega_t\}_{t=0}^{\infty}$ is assumed to be Markovian and the initial conditions $(k_0, \theta_0, \omega_0)$ are given[8].

Under the above constraints, the Second Welfare Theorem implies that Pareto Optimal (PO) allocations can be decentralized by a system of competitive markets for every given initial distribution of wealth. PO allocations

[8] A version of this problem was studied in Marcet and Marimon (1992). They had two agents, one risk averse and the other an –unconstrained– risk-neutral agent who acted as planner. The default value in that paper also depended on capital.

can be found by solving:

$$\max_{\{c_t, i_t\}} \quad \mathrm{E}_0 \sum_{t=0}^{\infty} \beta^t \sum_{j \in J} \alpha_j \; u(c_{j,t})$$

for positive weights α, subject to technological constraints and initial conditions. This problem is of the form of *Program 0* with $r(x,\ a,\ s) = \sum_{j\in J} \alpha_j \; u(c_j)$, $s \equiv (\theta, \omega)$; $x \equiv k$; $a \equiv (i, c)$; $\ell(x,\ a,\ s) \equiv (1-\delta)k + i$, and

$$A(x,s) \equiv \left\{ (i,c) \geq 0 \ : \ \sum_{j\in J} c_j + i \leq F(k,\theta) + \sum_{j\in J} \omega_j \right\} \ .$$

Therefore, standard dynamic programming is applicable, the usual Bellman equation is satisfied, and the optimal solution satisfies $(c_t, i_t) = f(k_t, \omega_t, \theta_t)$, where f is the decision function associated with the standard Bellman equation. Furthermore, if u is differentiable, optimal consumption allocations satisfy

$$\frac{u'(c_{i,t})}{u'(c_{j,t})} = \frac{\alpha_j}{\alpha_i}, \text{ for all } i, j \text{ and } t \tag{5}$$

In particular, when, by the First Welfare Theorem, the PO allocation is an *Arrow-Debreu* competitive allocation, $1/\alpha_j$ is agent j's marginal utility of income, which –in a strict form of the *Permanent Income Hypothesis*– remains constant through time, showing that individual consumption paths only depend on aggregate consumption and the initial wealth distribution.

The PO allocation can only be observed in economies where the planner has the ability to *enforce* the optimal contract by punishing *any* deviation from the optimal plan.[9] We now assume the enforcement technology available to the planner can not prevent any agent from switching to autarky in a given period and staying there forever. Then, the planner has to take into account the following participation constraints

$$\mathrm{E}_t \sum_{n=0}^{\infty} \beta^n \; u(c_{j,t+n}) \geq v_j^a(\omega_t) \qquad \text{for all } j, t \tag{6}$$

where $v_j^a(\omega_t) \equiv E_t \sum_{n=0}^{\infty} \beta^n \; u(\omega_{j,t+n})$.

[9]If the economy is decentralized by a system of competitive asset markets, full enforcement implies that all agents honor their debts.

Under these constraints, the planner's problem is of the form of *Program 1* since (6) is of the form (2) for $N = \infty$ once we let $g_2^j(x,\, a,\, s) \equiv u(c_j)$ and $g_1^j(x,\, a,\, s) \equiv u(c_j) - v_j^a(\omega)$.[10] Therefore, using the algebra in this subsection, it can be transformed into a *Program 2* with

$$H \equiv \mathrm{E}_0 \sum_{t=0}^{\infty} \beta^t \sum_{J} \left[(\alpha_j + \mu_{j,t})\ u_j(c_{j,t}) + \gamma_{j,t} \left(u(c_{j,t}) - v_j^a(\omega_t) \right) \right] \tag{7}$$

where

$$h(x,\, a, \mu,\, \gamma,\, s) \equiv \sum_{j \in J} \left((\alpha_j + \mu_j + \gamma_j) u(c_j) - \gamma_j\ v_j^a(\omega) \right).$$

Then, the solution of the SPP can be obtained by studying the **SPFE**

$$\begin{aligned} W(k,\, \mu, \omega, \theta) \ = \inf_{\gamma \geq 0} \sup_{c,i} \Bigg\{ & \sum_{j \in J} \left((\alpha_j + \mu_j')\ u(c_j) - \gamma_j\ v_j^a(\omega) \right) \\ & + \beta\ \mathrm{E}\left[W(k',\, \mu', \omega', \theta') \,|\omega, \theta \right] \Bigg\} \end{aligned} \tag{8}$$

$$\text{s.t. } k' = (1-\delta)k + i, \quad \sum_{j \in J} c_j + i \leq F(k, \theta) + \sum_{j \in J} \omega_j$$
$$\text{and } \mu' = \mu + \gamma$$

Letting ψ be the policy function associated with this functional equation, efficient allocations satisfy $(c_t,\, i_t, \gamma_t) \ = \ \psi(k_t,\, \mu_t,\, \theta_t,\, \omega_t)$ with initial conditions $(k_0,\, 0,\, \theta_0,\, \omega_0)$.

Notice that the objective function of (7) can be interpreted as if the weights that the planner assigns to each agent are shifting over time, according to whether or not the participation constraint is binding. Furthermore, if u is differentiable,

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

Thus, the optimal allocations amount to choosing efficiently the time profile of the time-dependent weights $(\alpha_j + \mu_{j,t+1})$, in such a way that the participation constraints are satisfied. Every time that the participation constraint for

[10] Clearly, the function v^a can be found without knowledge of the solution.

an agent is binding, his weight is increased by the amount of the corresponding lagrange multiplier. An agent is induced not to default by increasing his consumption not only in the period where he is tempted to default, but also for many of the following periods; in this way, the additional consumption that the agent receives to prevent default is smoothed over time. That is, individual paths of consumption depend on individual histories (in particular, on past "temptations to default") not just on the initial wealth distribution and the aggregate consumption path, as in the *Arrow-Debreu* competitive allocations. This also shows that if enforcement constraints are never binding (e.g., punishments are severe enough) then $\mu_t = \mu_0$ and we recover the "constancy of the marginal utility of expenditure". In other words, the evolution of the co-state variables can be also interpreted as the evolution of the distribution of wealth[11].

In Section 3 we state an interiority condition that is needed for existence of a SPP; also, some convexity conditions are needed for uniqueness of the solution. These conditions are trivially satisfied in this example under standard strict concavity assumptions.

2.2 Two-period intertemporal constraints

Consider now the case where $N = 1$ in (2). Intertemporal constraints of this form arise in dynamic Stackelberg games. For example, in dynamic Ramsey problems, where the government chooses policy variables subject to the Euler equations satisfied in equilibrium. Example 2 below is one of these cases.

Let $\beta^{-t}\gamma_t$ be the Lagrange multiplier of (2). We proceed as in Subsection 2.1, constructing the Lagrangean, applying the law of iterated expectations and reordering terms in order to group together terms that depend on information available at t. We can check with simple algebra that the objective function for the Lagrangian of this problem takes the form

$$H = \mathrm{E}_0 \sum_{t=0}^{\infty} \beta^t \left(r(x_t, a_t, s_t) + \gamma_t g_1(x_t, a_t, s_t) + \mu_t g_2(x_t, a_t, s_t) \right) \qquad (9)$$

for all feasible sequences, letting $\mu_t = \gamma_{t-1}$ and $\mu_0 = 0$.

The saddle point of H is of the form of *Problem 2*, taking

[11]Kletzer and Wright (1998) apply an example, similar to the one presented here, to discuss sovereign debt problems.

$$\begin{aligned} h(x,\, a, \mu,\, \gamma,\, s) &\equiv h_0(x,\, a,\, s) + \gamma h_1(x,\, a,\, s) + \mu h_2(x,\, a,\, s) \\ &\equiv r(x,\, a,\, s) + \gamma g_1(x,\, a,\, s) + \mu g_2(x,\, a,\, s) \end{aligned}$$

$$\varphi(\mu,\, \gamma,\, s) \equiv \gamma, \qquad\qquad \mu_0 = 0$$

Example 2. A Ramsey problem

We consider a simple model of Ramsey equilibrium, which has a constraint of the form (2) for $N = 1$. There is a constant returns to scale technology that, when labor is normalized to one, reduces to the technology of Example 1. There is a representative agent who rents capital to a firm and inelastically supplies one unit of labor. Capital and labor markets are competitive, but no financial assets are available to the government. Government spending $\{c_t^g\}$ generates utility for the consumer, and it is financed by levying an income tax, τ_t, that equally taxes capital and labor income. Furthermore, government's budget must be balanced in each period. The problem of optimal policy consists of choosing among different combinations of government spending and taxes that satisfy the budget constraint of the government and that are compatible with competitive equilibrium.

This example adds one complication relative to previous papers on optimal taxation since, even using the primal approach, the implementability constraints can not be summarized in one constraint as is done, for example, in Lucas and Stokey (1983) and Chari, *et al.* (1995). This is due to the absence of financial assets. Other than this, our example is meant to be as simple as possible.

The representative consumer solves the problem

$$\begin{aligned} &\max_{\{c_t, i_t\}} \quad \mathrm{E}_0 \sum_{t=0}^{\infty} \beta^t \;\; [u(c_t) + v(c_t^g)] \\ &\qquad\qquad \text{s.t.} c_t + k_{t+1} \leq r_t k_t (1 - \tau_t) + (1 - \delta) k_t + w_t (1 - \tau_t) \end{aligned} \tag{10}$$

The Euler, first order, condition for the consumer is

$$u'(c_t) = \beta \mathrm{E}_t \left[u'(c_{t+1}) \; (r_{t+1}(1 - \tau_{t+1}) + 1 - \delta) \right] \tag{11}$$

Since prices are competitive, $r_t = F'(k_t, \theta_t)$ and $w_t = F(k_t, \theta_t) - r_t k_t$. The budget constraint of the government in period t is,

$$c_t^g = \tau_t(r_t \; k_t + w_t) \; = \; \tau_t \; F(k_t, \theta_t) \tag{12}$$

Following Ramsey's principle of optimal taxation, the government maximizes the representative consumer's tastes subject to feasibility constraints and the implementability constraints (11) and (12).

Strictly speaking, the 'equal' sign in the above Euler equation means that the set of allocations that is feasible to the planner has an empty interior. In the next section it is shown that, as usual, the Lagragean approach is sufficient for a solution only if an interior point exists. In appendix 3 we show that, in the usual case, the solution is equivalent to a problem where the Euler equation is written as a weak inequality. Combining all these observations, we see that the Ramsey problem maximizes the utility of the agent subject to the *implementability* constraint

$$u'(c_t) \leq \beta E_t \left[u'(c_{t+1}) \left(F'(k_{t+1}, \theta_{t+1}) \left(1 - \frac{c_{t+1}^g}{F(k_{t+1}, \theta_{t+1})} \right) + 1 - \delta \right) \right].$$

Then, the Ramsey problem is of the form of *Program 1* with $s \equiv \theta$; $x = k$; $a = (i, c, c^g)$, $r(x, a, s) = u(c) + v(c^g)$, $\ell(x, a, s) \equiv (1-\delta)k + i$, $A(x, s) = \{(i, c, c^g) \geq 0 : i + c + c^g \leq F(k, \theta)\}$ and the constraints (2) given by

$$\begin{aligned} g_1(x, a, s) \; &\equiv \; -u'(c) \\ g_2(x, a, s) \; &\equiv \; u'(c) \; \left(\; F'(k, \theta) \; \left(1 - \frac{c^g}{F(k, \theta)} \right) + 1 - \delta \right). \end{aligned}$$

Then, the SPP of *Program 2* holds for

$$\begin{aligned} h(x, a, \mu, \gamma, s) \; &\equiv \; u(c) + v(c^g) + \gamma \; u'(c) \\ & \quad -\mu \; u'(c) \; \left(F'(k, \theta) \left(1 - \frac{c^g}{F(k, \theta)} \right) + 1 - \delta \right) \end{aligned} \tag{13}$$

The corresponding **SPFE** is

$$\begin{aligned} W(k, \mu, \theta) \;\; &= \;\; \inf_{\gamma \geq 0} \sup_{c, i, \mathbf{c^g}} \{ h(x, a, \mu, \gamma, s) + \beta \; \mathrm{E}[W(k', \mu', \theta') \,|\theta]\} \\ & \quad \text{s.t. } k' = (1-\delta)k + i, \quad c + i + c^g \leq F(k, \theta) \\ & \quad \text{and } \mu' = \gamma \end{aligned}$$

With the policy function ψ associated to this **SPFE**, we can find the optimal solution to the Ramsey problem by setting $(c_t, i_t, c_t^g, \gamma_t) = \psi(k_t, \theta_t, \mu_t)$ for all t, with $\mu_0 = 0$. This uniquely defines a stationary policy for τ in an obvious way.

As in the previous example, sufficient conditions for the interiority and convexity conditions are discussed in appendix 3. The interiority condition is satisfied in most applications. The convexity conditions can only be obtained under restrictive assumptions; obviously, this is not a problem with our approach, but a problem that often arises in Ramsey equilibria.

2.3 Other applications and relation to the literature

A number of numerical algorithms can be used to compute (or approximate) the function ψ. A number of applications are already present in the literature, and they have used different algorithms. For example, Marcet and Marimon (1992), Marcet, *et al.* (1996) and Rojas (1993) approximate the first order conditions with PEA, Kehoe and Perry (1998) perform backward iterations on the value function, which is related to iterating the above **SPFE**, and Hansen, *et al.* (1985) compute the solution in the linear case using traditional tools of linear dynamic programming. In a follow up paper we provide several examples and the details in computing solutions to those examples.

Other approaches are available in the literature trying to reduce the dimension of history-dependent optimal contracts (or equilibrium strategies). We briefly mention and discuss those approaches. A detailed account of advantages and disadvantages of different methods can only be done in the context of concrete examples and it falls beyond the scope of this paper.

The work of Hansen, *et al.* (1985) can be seen as an early application of the approach we take in this paper. They formulated the problem with lagrange multipliers as co-state variables in a linear model with a two-period constraint like the one discussed in section 2.2. They also discussed the relation of the fixed initial condition $\mu_0 = 0$ with the time-inconsistency problem. In another piece of early work, Kydland and Prescott (1980) propose to include the lagrange multiplier of the budget constraint of the consumer as a co-state variable[12].

[12]Our work provides a formal proof that introducing the co-state variables is sufficient for the optimal solution in Hansen, *et al.* (1985). It is not clear if the approach of Kydland and Prescott (1980) does provide an optimal solution in cases with uncertainty.

The primal approach, as applied in Lucas and Stokey (1983) and Chari, *et al.* (1995), can be used in many problems of optimal policy. If the government can complete the markets with its policy instruments, all the implementability constraints are summarized into one, and the vector of state variables is not enlarged relative to the standard case, after period 1. Within our framework, this means that the co-state variables are constant; in most models (such as examples 1 and 2 and the applications just cited), the co-state variables need to be introduced[13].

Rustichini (1996) discusses a model similar to our example 1 and he artificially *imposes* the restriction that the solution depends only the natural state variables, in that case (k, ω, θ). He defines a map that delivers a value function satisfying incentive constraints under this restriction. In most models, the map defined by this author does not reach the optimum under full commitment, while our approach would. In those models, it is precisely the introduction of the co-state variables that allows consumption smoothing across periods (see our discussion of consumption smoothing at the end of example 1); the solution computed by the approach of Rustichini (1996) would not allow to spread the compensations for default over different periods, but agents receive a single one-shot compensation when they are tempted to default. Thus the optimum under full commitment is not achieved.

The pioneer work of Abreu, *et al.* (1990) proposes to summarize past histories in the *function* of promised utilities. This amounts to using a *function* as a co-state variable, which still leaves for a fairly large state space. In some cases, (see Green (1987), and Thomas and Worral (1990)) this *function* can be summarized into a few co-state variables[14]. This approach is used by Phelan and Townsend (1992). Our approach provides a common framework that encompasses two-period constraints as well as discounted sums other than discounted utilities (as is the case with some present value budget constraints); it makes it possible to directly obtain and characterize efficient

[13]See, for example, Benhabib, *et al.* (1997) for an "optimal" tax policy derived using non-constant co-state variables.

[14]See also Chang (1996) for a general disscussion, and an application to the design of credible monetary policies, of the APS approach, as well as the similar approach pioneered by Cronshaw and Luenberger (1994). See, for example, Sargent and Ljungqvist (1998) and Sargent (1999) for some macroeconomic applications of the approach of using first backwards iteration of "future present values" (i.e., APS) and then maximizing over the set of feasible –incentive compatible– contracts.

contracts on the Pareto frontier without having to characterize the whole set of possible equilibria and, ultimately, as we discussed in the introduction, to obtain fully recursive solutions, in the sense that a time-invariant map is defined *and* initial conditions are given. In contrast, using the APS approach, the initial value function at time zero is not known and needs to be solved for separately. Having to solve for the initial condition may lead to unstable outcomes, much in the same way as with solving Euler equations by forward shooting. Nevertheless, the APS approach is more suitable to characterize the set of all incentive compatible contracts (i.e., sub-game perfect equilibria) and (until the theory in this paper is not further developed) to study problems with informational constraints.

3 Programs 1 and 2 and their duality

In this Section we study first *Program 1* and show how, under suitable assumptions, it has solutions, which have a Lagrangean formulation. The underlying theory is well known (see, for example, Luenberger (1969)) and we extend it to cover constraints of the form (2). We then, study *Program 2* and show that, under appropriate conditions, there exist *saddle-point* solutions. Treating *Program 2* as a primitive program allows us to develop a recursive theory that does not specifically rely on *Program 1*. Nevertheless, we show how to derive *Program 2* from *Program 1*. We end the section with a global *duality* theorem linking both programs (which requires convexity assumptions) and a *sufficiency* theorem showing that *Program 2* solutions are *Program 1* solutions (with weaker convexity requirements). We call *Program 1* the Planner's Problem (**PP**) and *Program 2* the Saddle Point Problem (**SPP**).

3.1 Program 1 (PP)

We first recall *Program 1*

PP

$$V(x_0,\, s_0) = \sup_{\{a_t\}} \quad \mathrm{E}_0 \sum_{t=0}^{\infty} \beta^t r(x_t,\, a_t,\, s_t)$$

$$\text{s.t.} \quad x_{t+1} = \ell(x_t,\, a_t,\, s_{t+1}),\;\; a_t \in A(x_t,\, s_t),\;\; t \geq 0 \qquad (1a)$$

$$g_1^j(x_t,\, a_t,\, s_t) \;\; + \;\; \mathrm{E}_t \left[\sum_{n=1}^{N_j} \beta^n g_2^j(x_{t+n},\, a_{t+n},\, s_{t+n}) \right] \geq 0;\; j = 1, ..., k \qquad (2)$$

$$x_0,\, s_0 \text{ given.}$$

$$a_t \qquad \mathcal{S}_t - \text{measurable}$$

We make the following assumptions[15],

A1. S is a compact (Borel) set of an Euclidean space. $\{s_t\}$,$s_t \in S$, is a Markovian process satisfying the following (Feller) property: if $f : S \to \mathcal{R}$ is bounded and continuous (i.e., $f \in C(S)$), then $E[f|\cdot] : S \to \mathcal{R}$ is also bounded and continuous (i.e., $E[f|\cdot] \in C(S)$)[16].

A2. X is a compact subset of $\mathcal{R}^\ell$; $A(\cdot,\, s)$ is a compact, convex valued and continuous correspondence from X to $\mathcal{R}^m$. Furthermore, $\ell(\cdot, \cdot\,,\, s) : X \times \mathcal{R}^m \to X$ is continuous.[17]

A3. $r(\cdot,\, \cdot, s) : \mathcal{R}^\ell \times \mathcal{R}^m \to \mathcal{R}$ is continuous and bounded. Furthermore, $\beta \in (0, 1)$.

A4. $g_n^j(\cdot,\, \cdot,\, s) : \mathcal{R}^\ell \times \mathcal{R}^m \to \mathcal{R}, n = 1, 2, j = 1, ..., k$ is continuous and bounded.

A4b. $g_n^j(\cdot,\, \cdot,\, s), n = 1, 2,\, j = 1, ..., k$ and $r(\cdot,\, \cdot, s)$ are quasiconcave, and the set of $\{a_t\}$ satisfying $(1a)$ is convex [18].

A5. There exists an $\epsilon > 0$ and, for all $(x_0,\, s_0)$, a program $\{\hat{a}_n\}$ satisfying $(1a)$ such that $d(\hat{a}_n, A(\widehat{x}_n,\, s_n)^c) \geq \varepsilon$ and[19]

$$g_1^j(x_0,\, \hat{a}_0,\, s_0) + \mathrm{E}_t \left[\sum_{n=1}^{N_j} \beta^n g_2^j(\hat{x}_n,\, \hat{a}_n,\, s_n) \right] \geq \epsilon$$

---

[15] We denote by $\mathcal{S}_t$ the σ-field generated by all possible sequences $(s_0, ..., s_t)$.

[16] If the underlying transition kernel of $\{s_t\}$ is Q, then $E[f|s] \equiv \int f(s')Q(s, ds')$.

[17] Our assumptions, such as **A2**, need only to be satisfied "for almost all $s \in S$", with respect to the probabilities defined by the transition probability; i.e., $Q(s, \cdot)$.

[18] This last condition is satiafied if, for example, $A(\ell(x, \cdot\,,\, s),\, s)$ has a convex graph for all (x, s).

[19] As usual, the superindex ^{c} on a set denotes the complement of this set; d denotes the Euclidian distance between a point and a set.

where $\hat{x}_{n+1} = \ell(\hat{x}_n,\, \hat{a}_n,\, s_{n+1})$ and, for each $j = 1, ..., k$, either $N_j = \infty$ or $N_j = 1$.

Assumptions **A1** - **A3** are standard in stochastic maximization problems. Assumption **A5** is a generalized version of a standard interiority assumption. Assumption **A4** corresponds to the new set of constraints (2). It requires that the constraint set is well behaved (closed and bounded) even when these constraints are taken into account. The convexity assumption **A4b** is common in maximization problems, but fairly restrictive in many applications with implementability constraints. As we will see, this assumption is only required to obtain global duality results that guarantee the equivalence between solutions of PP and solutions of SPP, but is not needed for many of our results[20]. In particular, *Lagrangean* multipliers may exist and value functions may be well defined even **A4b** is not satisfied (see, for example, Luenberger (1969) and Stokey, *et al.* (1989)). In that case, solutions of a Lagrangean are also solutions to the maximization problem.

The infinite-dimensional formulation of PP

We can describe, more compactly, **PP** as a maximal problem in $\mathcal{L}_\infty$. Given an initial condition (x_0, s_0), let

$$\begin{array}{c} f(\mathbf{a}) = \mathrm{E}_0 \sum_{t=0}^{\infty} \beta^t r(x_t, a_t, s_t) \\ \text{s.t. } \; x_t = \ell(x_{t-1},\, a_{t-1},\, s_t) \text{ for } t > 0 \end{array}$$

and let $g(\mathbf{a})$ be defined, coordinatewise, as

$$g(\mathbf{a})_t = \left[\begin{array}{c} d(a_t, A(x_t\,,\, s_t)^c) \\ g_1^1(x_t,\, a_t,\, s_t) + \mathrm{E}_t \left[\sum_{n=1}^{N_1} \beta^n g_2^1(x_{t+n},\, a_{t+n},\, s_{t+n}) \right] \\ \cdots \\ g_1^k(x_t,\, a_t,\, s_t) + \mathrm{E}_t \left[\sum_{n=1}^{N_k} \beta^n g_2^k(x_{t+n},\, a_{t+n},\, s_{t+n}) \right] \end{array} \right]$$
$$\text{s.t. for } t > 0,\; x_t = \ell(x_{t-1},\, a_{t-1},\, s_t)$$

With this notation, **PP** with initial condition $(x_0,\, s_0)$ becomes

[20] In particular, unless it is explicitly mentioned, we will not make such assumption; e.g., **A1-A5** means **A1,A2,A3,A4 and A5**.

$$\textbf{(PP)} \qquad \sup_{g(\mathbf{a})\geq 0} f(\mathbf{a})$$

This is a standard maximization problem in $\mathcal{L}_\infty$, the following proposition shows that our assumptions guarantee the existence of a solution to **(PP)**.

Proposition 1 *Assume* **A1-A3**, **A4** *and* **A5**. *There exists a program* $\mathbf{a}^*$ *which solves* **PP** *with initial condition* (x_0, s_0).

Before proving the first proposition, it is convenient to be more explicit with respect to the underlying commodity space. By **A1**, there is a well defined probability space $(S_\infty, \mathcal{S}_\infty, P)$. If $\mathcal{L}^m_\infty(S_\infty, \mathcal{S}_t, P)$ denotes the space of m-valued –essentially bounded– $\mathcal{S}_t$-measurable functions, then the *contract space* is $\mathcal{A} = \{\mathbf{a} : \forall t \geq 0,\ a_t \in \mathcal{L}^m_\infty(S_\infty, \mathcal{S}_t, P)\}$. The *contract space*, $\mathcal{A}$, has a product structure. We consider the *product topology*, characterized by having every projection –say, to $\mathcal{L}^m_\infty(S_\infty, \mathcal{S}_t, P)$– with its corresponding topology. For example, when we refer to the weak*-topology, we mean that such topology is considered in any projection (i.e., the $\sigma(\mathcal{L}_\infty, \mathcal{L}_1)$ topology on $\mathcal{L}_\infty(S_\infty, \mathcal{S}_t, P)$)[21]. We will make use of the *topology of convergence in probability*, denoted P-topology. [22] An important feature of the P-topology is that, with this topology, $\mathcal{L}_\infty(S_\infty, \mathcal{S}_t, P)$ is a complete metric space[23].

Proof: By assumption **A5** there are feasible solutions. Given assumptions **A1, A2** and **A4**, it is easy to see that the set $\{\mathbf{a} \in \mathcal{A} : g(\mathbf{a}) \geq 0\}$ is closed and *totally bounded* with respect to the *P-topology* (i.e., for any $\epsilon > 0$, it can be covered with a finite number of spheres of radius ϵ). Since the *P-topology* defines a complete metric space, it follows (see, Dunford and Schwartz (1957), Theorem I.6.15, p.22) that $\{\mathbf{a} \in \mathcal{A} :$

[21]Our approach extends to the more general case where, for every $t > 0$, $a_t \in \mathcal{A}_t$ and $\mathcal{A}_t$ is an arbitrary linear space, provided that **A1-A5** are appropriately modified as to guarantee that a solution to **PP** (or **SPP**) exists. Our approach relies on the time separability of the objective function and of the constraint sets (or some weaker time recursive form of these maps) and, therefore, does not depend on the specific topological structure of $\mathcal{L}^m_\infty$.

[22]A metric d_p can be defined on $\mathcal{L}_\infty(S_\infty, \mathcal{S}_t, P)$ by $d_p(x, \hat{x}) = \int \frac{|x-\hat{x}|}{1+|x-\hat{x}|} P(ds)$. This metric induces the topology of convergence in probability.

[23]See, for example, (Neveu (1970), Pr. II 3-4).

$g(\mathbf{a}) \geq 0\}$ is compact. By assumptions **A2** & **A3** $r(\cdot, \cdot, \cdot)$ is uniformly integrable, therefore $f(\cdot)$ is continuous with respect to the *P-topology*. It follows that a maximal element, $\mathbf{a}^*$, exists.■

This proposition shows that **PP** has a well defined value. However, more than in the existence of solutions, we are interested in their characterization and, in particular, in attaining a recursive formulation. To this end, we study first the *Lagrangean* structure of **PP**. An interesting feature of **PP** is that g: $\mathcal{A} \to \mathcal{L}_\infty$ and, as it is well known, the positive orthant of $\mathcal{L}_\infty$ has a non empty interior, as it is required for a standard separation argument[24].

Proposition 2 Assume **A1**- **A4, A4b** & **A5**. Let $\mathbf{a}^*$ be a solution to **PP** with initial condition $(x_0,\, s_0)$. There exist a $\tilde{\gamma}^* \in \mathcal{L}_1$ such that the *lagrangean*

$$L(\mathbf{a},\, \tilde{\gamma}) = f(\mathbf{a}) + \tilde{\gamma} g(\mathbf{a})$$

has a **saddle point** at $(\mathbf{a}^*,\, \tilde{\gamma}^*)$, i.e.,

$$L(\mathbf{a}^*,\, \tilde{\gamma}) \geq L(\mathbf{a}^*,\, \tilde{\gamma}^*) \geq L(\mathbf{a},\, \tilde{\gamma}^*) \tag{14}$$

for all $\mathbf{a} \in \mathcal{A}$ and $\tilde{\gamma} \in \mathcal{L}_{1,+}$. Furthermore, $V(x_0,\, s_0) = L(\mathbf{a}^*,\, \tilde{\gamma}^*)$.

Proof See Appendix 1.

3.2 Program 2 (SPP)

In this subsection we take *Program 2* as our primitive program and show the existence of solutions. *Program 2* is a *Saddle Point Problem* (**SPP**) and, given arbitrary initial conditions $(x_0,\, \mu_0,\, s_0)$ (i.e., in this Section we do not constrain $\mu_0 = 0$), its value –possibly, infinity– is $W(x_0,\, \mu_0,\, s_0)$:

SPP

[24] Even if a standard separation argument can not be applied, because the assumptions underlying **PP** are not satisfied, it may be possible to extend our approach, as long as *Lagrange multipliers* are well defined (and summable). This can be the case when $g(\cdot)$ maps into a more general space –say, using separation arguments as in Mas-Colell and Zame (1991) – or when the problem is not a global convex problem and *Lagrange multipliers* are for example derived as dual variables in a smooth local optimization problem (see the Remark following Theorem 2).

$$
\begin{aligned}
W(x_0,\, \mu_0,\, s_0) = \inf_{\{\gamma_t\}} \sup_{\{a_t\}} \quad & \left\{ \mathrm{E}_0 \sum_{t=0}^{\infty} \beta^t h(x_t,\, a_t,\, \mu_t,\, \gamma_t,\, s_t) \right\} \\
s.t. \quad & x_{t+1} = \ell(x_t,\, a_t,\, s_{t+1}), \quad a_t \in A(x_t,\, s_t),\ t \geq 0 \qquad (1a) \\
& \mu_{t+1} = \varphi(\mu_t,\, \gamma_t,\, s_{t+1}), \quad \gamma_t \geq 0,\ t \geq 0 \qquad (3) \\
& (x_0, \mu_0, s_0) \ \textit{given.} \\
& (a_t, \gamma_t), \mathcal{S}_t - \textit{measurable}
\end{aligned}
$$

We make the following assumptions, in addition to **A1**-**A2,**

B1. $h(\cdot,\, \cdot,\, \cdot \cdot,\, s)$ is continuous; $h(\cdot,\, \cdot,\, \mu,\, \gamma,\, s)$ is bounded and h is of the form

$$h(x,\, a, \mu,\, \gamma,\, s) \equiv h_0(x,\, a,\, s) + \gamma h_1(x,\, a,\, s) + \mu h_2(x,\, a,\, s)$$

Furthermore, $\beta \in (0, 1)$[25]

B1b. $h(\cdot,\, \cdot,\, \cdot,\, \cdot,\, s)$ is quasiconcave, and the set of $\{a_t\}$ satisfying $(1a)$ is convex .

B2. There exists an $\epsilon > 0$ and, for all $(x_0,\, s_0)$, a program $\{\hat{a}_n\}$ satisfying $(1a)$ with $d(\hat{a}_n, A(\widehat{x}_n,\, s_n)^c) \geq \varepsilon$ such that,

$$h_1^j(x_0,\, \hat{a}_0,\, s_0) + \mathrm{E}_0 \left[\sum_{n=1}^{N_j} \beta^t h_2^j(\hat{x}_n,\, \hat{a}_n,\, s_n) \right] \geq \epsilon$$

where $\hat{x}_{n+1} = \ell(\hat{x}_n,\, \hat{a}_n,\, s_{n+1})$. For $j = 1, ..., k$, either $N_j = \infty$, in which case $\varphi_j(\mu,\, \gamma,\, s) = \mu_j + \gamma_j$, or $N_j = 1$, in which case $\varphi_j(\mu,\, \gamma,\, s) = \gamma_j$.

As we will see in the next subsection (and should be clear from Section 2), when *Program 2* is obtained from *Program 1*, assumptions **B1**-**B2** are satisfied whenever *Program 1* satisfies **A3**-**A5**.

Given an initial condition $(x_{0,}\mu_{0,}s_0)$, it is also possible to write **SPP** in a more compact form by letting

$$
\begin{aligned}
H(\mathbf{a}, \boldsymbol{\gamma}) \ &= \mathrm{E}_0 \textstyle\sum_{t=0}^{\infty} \beta^t h(x_t,\, a_t,\, \mu_t,\, \gamma_t,\, s_t) \\
\text{s.t. for } t > 0,\ x_t &= \ell(x_{t-1},\, a_{t-1},\, s_t) \text{ and } \mu_t = \varphi(\mu_{t-1},\, \gamma_{t-1},\, s_t)
\end{aligned}
$$

[25] We use the notation $\mu h_2(x, a,\, s)$ to denote $\sum_{j=1}^{k} \mu^j [h_2(x,\, a,\, s)]^j$.

$$q(\mathbf{a})_t = d(a_t, A(x_t, s_t)^c)$$
$$s.t.\ for\ t > 0,\ x_t = \ell(x_{t-1}, a_{t-1}, s_t)$$

With this notation, the saddle point problem takes the form

$$W(x_0, \mu_0, s_0) = \inf_{\boldsymbol{\gamma} \geq 0} \sup_{q(\mathbf{a}) \geq 0} H(\mathbf{a}, \boldsymbol{\gamma})$$

We will also use the value function decomposition

$$\begin{aligned} H(\mathbf{a}, \boldsymbol{\gamma}) &\equiv H_0(\mathbf{a}) + H_1(\mathbf{a}, \boldsymbol{\gamma}) \\ &\equiv \mathrm{E}_0 \sum_{t=0}^{\infty} \beta^t \left[h_0(x_t, a_t, s_t)\right] + \mathrm{E}_0 \sum_{t=0}^{\infty} \beta^t \left[\gamma_t h_1(x_t, a_t, s_t) + \mu_t h_2(x_t, a_t, s_t)\right] \end{aligned}$$

which leads to the following decomposition

$$W(x, \mu, s) \equiv W_0(x, \mu, s) + W_1(x, \mu, s)$$

When **SPP** is derived from **PP,** then Propositions 1 and 2 guarantee the existence of solutions to **SPP**. Nevertheless, we are interested in treating **SPP** as a primitive problem. The following proposition shows that, with assumptions **A1-A2** and **B1-B1b-B2**, **SPP** has a solution and the existence result follows from a fixed point theorem. As in the existence of Nash equilibria, convexity is a necessary condition for the application of -a generalized version of- Kakutani's fixed point theorem, but, as in games, solutions may exist even when convexity fails (see Appendix 1 for its proof, as well as for the proof of its Corollary)[26].

Proposition 3 *Assume* **A1-A2** *and* **B1-B1b-B2**. *Given initial conditions* (x_0, μ_0, s_0)*, there exists a solution* $(\mathbf{a}^*, \boldsymbol{\gamma}^*)$ *to* **SPP.** *Furthermore, all solutions to* **SPP** *have value* $W(x_0, \mu_0, s_0)$.

The following corollary provides bounds on multipliers and it is of particular interest for developing a recursive formulations, as well as computational solutions. Assumption **B2** plays a key role, while it does not rely on the convexity assumption **B1b.**

[26]We let $\|\boldsymbol{\gamma}\|_\beta = \sum_{t=0}^{\infty} \beta^t \|\gamma_t\|$.

Corollary to Proposition 3. Assume **A1-A2** and **B1-B2**. Let $(\mathbf{a}^*, \boldsymbol{\gamma}^*)$ be a solution to **SPP.** There exist a positive constant $\bar{K}$ such that $(\mathbf{a}^*, \boldsymbol{\gamma}^*)$ satisfies $\|\boldsymbol{\gamma}^*\|_\beta < \bar{K} \cdot \max\{1, \|\mu_0\|\}$.

3.3 Duality

We finish this Section by relating *Program 1* and *Program 2.* Theorem 1 provides a *duality theorem for recursive contracts* in which we derive **SPP** from **PP** using Proposition 2, which requires the convexity assumption A4b. Theorem 2 is a *sufficiency theorem for recursive contracts*: SPP solutions are PP solutions. This second theorem exploits the fact that SPP can be derived from PP even when the global convexity assumptions (of Theorem 1) are not satisfied (as long as summable multipliers exist).

We now make it clear how to derive SPP from PP. To this end, notice that $\tilde{\boldsymbol{\gamma}} g(\mathbf{a}) = \mathbf{E}_0 \sum_{t=0}^{\infty} \sum_{j=0}^{k} \tilde{\gamma}_t^j g^j(\mathbf{a})_t$, and that, for $j = 1, ..., k$,

$$\begin{aligned} \mathrm{E}_0 \sum_{t=0}^{\infty} \tilde{\gamma}_t^j g_t^j(\mathbf{a}) &= \mathrm{E}_0 \sum_{t=0}^{\infty} \tilde{\gamma}_t^j \left[g_1^j(x_t, a_t, s_t) + \mathrm{E}_t \sum_{n=1}^{N_j} \beta^n g_2^j(x_{t+n}, a_{t+n}, s_{t+n}) \right] \\ &= \mathrm{E}_0 \sum_{t=0}^{\infty} \beta^t \gamma_t^j \left[h_1^j(x_t, a_t, s_t) + \sum_{n=1}^{N_j} \beta^n h_2^j(x_{t+n}, a_{t+n}, s_{t+n}) \right] \\ &\equiv \mathrm{E}_0 \sum_{t=0}^{\infty} \beta^t \left[\gamma_t^j h_1^j(x_t, a_t, s_t) + \mu_{t+1}^j h_2^j(x_{t+n}, a_{t+n}, s_{t+n}) \right] \end{aligned}$$

where the second equality follows from identifying $h_m^j(\cdot)$ with $g_m^j(\cdot)$, $m = 1, 2, \gamma_t^j$ with $\tilde{\gamma}_t^j \beta^{-t}$ and applying the law of iterated expectations. The third follows from simple algebra and taking $\varphi(\mu, \gamma, s) = \mu + \gamma$ if $N_j = \infty$ and $\varphi(\mu, \gamma, s) = \gamma$ if $N_j = 1$. Now identify $h_0^j(\cdot)$ with $r(\cdot)$ and let $\mu_0 = 0$, then it follows that $H_0(\mathbf{a}) = f(\mathbf{a})$ and $H_1(\mathbf{a}, \boldsymbol{\gamma}) + \tilde{\boldsymbol{\gamma}}^0 q(\mathbf{a}) = \tilde{\boldsymbol{\gamma}} g(\mathbf{a})$, where $\tilde{\gamma}_t^0$ is the lagrange multiplier associated with the resource constraint $q(\mathbf{a})_t \geq 0$. This shows how an **SPP** satisfying **B1**-**B1b**-**B2** can be derived from **PP.**

Theorem 1. Let **PP**, with initial condition (x_0, s_0), satisfy **A1–A4, A4b & A5**. The **SPP** (derived from **PP)** with initial condition $(x_0, 0, s_0)$ satisfies **A1–A2** and **B1**-**B1b-B2**. Furthermore, $\mathbf{a}^*$ is a solution to

PP if and only if there exist $\boldsymbol{\gamma}^*$ and $(\mathbf{a}^*, \boldsymbol{\gamma}^*)$ is a solution to **SPP** (derived from **PP**); that is, $W(x_0, 0, s_0) = V(x_0, s_0)$.

Proof : The proof consist on deriving **PP** from **SPP**, assembling the results of the last two subsections and applying *Lagrange Duality* theory (see, for example, Luenberger (1969), 8.6, Theorem 1). In particular, the theorem follows from four basic facts:

1. $\mathbf{a}^*$ is a solution to **PP**, with initial condition (x_0, s_0), if and only if $\mathbf{a}^*$ is a solution of $\sup_{g(\mathbf{a})\geq 0} f(\mathbf{a})$ (when f and g are defined with respect to the initial condition (x_0, s_0)).

2. $\max_{g(\mathbf{a})\geq 0} f(\mathbf{a}) = \min_{\tilde{\gamma}\in\mathcal{L}_{1,+}} \max_{\mathbf{a}\in\mathcal{A}} [f(\mathbf{a}) + \tilde{\gamma} g(\mathbf{a})]$

3. $\min_{\tilde{\gamma}\in\mathcal{L}_{1,+}} \max_{\mathbf{a}\in\mathcal{A}} [f(\mathbf{a}) + \tilde{\gamma} g(\mathbf{a})] = \min_{\gamma\geq 0} \max_{q(\mathbf{a})\geq 0} H(\mathbf{a}, \gamma)$

4. $(\mathbf{a}^*, \boldsymbol{\gamma}^*)$ is a solution to **SPP**, with initial condition $(x_0, 0, s_0)$, if and only if $(\mathbf{a}^*, \boldsymbol{\gamma}^*)$ is a solution of $\inf_{\gamma\geq 0} \sup_{q(\mathbf{a})\geq 0} H(\mathbf{a}, \gamma)$

The first and the fourth fact are immediate from our constructions in the last two subsections. The second fact is the *Lagrange Duality* result, when a solution $\mathbf{a}^*$ exists (which it does by Proposition 1).

To see the third fact notice that, by Proposition 2, when $(\mathbf{a}^*, \boldsymbol{\gamma}^*)$ is a solution to **SPP** and $\boldsymbol{\beta\gamma}^* \equiv (\gamma_0^*, \beta\gamma_1^*, ..., \beta^t\gamma_t^*, ...)$, then $\boldsymbol{\beta\gamma}^* \in \mathcal{L}_{1,+}$. Furthermore, given the convexity and interiority assumptions on resource constraints one can also show that, for the maximization part of $H(\mathbf{a}, \boldsymbol{\gamma}^*)$, there is a multiplier associated with such constraint (a version of Proposition 2 for the max part of **SPP**). That is, $H(\mathbf{a}, \gamma) + \tilde{\gamma}^0 q(\mathbf{a})$ defines a Lagrangian for **SPP**. It follows that, with the previous identification of maps and multipliers, it is also possible to go from **SPP** to the Lagrange formulation of **PP**.■

3.4 Sufficiency

Theorem 1 assumes convexity (i.e., **A4b**) to guarantee that there exist an appropriate multiplier, however, as long as **SPP** has a solution then it is a solution to the PP problem. Proposition 3 assumes convexity (i.e., **B1b)** *to* guarantee the existence of a solution to **SPP,** however, it is not a necessary condition for existence of solutions to **SPP.** The following sufficiency theorem does not rely in any convexity assumption and, therefore, can be a starting point for problems whose **SPP** formulation is known to have a solution.

Theorem 2. Given initial conditions $(x_0, 0, s_0)$, let $(\mathbf{a}^*, \boldsymbol{\gamma}^*)$ be a solution to **SPP**. Then $\mathbf{a}^*$ is a solution to **PP.** Furthermore, $W(x_0, 0, s_0) = V(x_0, s_0)$.

Proof: The proof is an extension, to **SPP**, of a sufficiency theorem for Lagrangian saddle points (see, for example, Luenberger (1969), Theorem 8.5.2, p.221). First notice that since $\mu_0 = 0$ (and with the simple algebra used in the proof of Theorem 1)

$$\begin{aligned} H(\mathbf{a}, \boldsymbol{\gamma}) \;&\equiv\; H_0(\mathbf{a}) + H_1(\mathbf{a}, \boldsymbol{\gamma}) \equiv \mathrm{E}_0 \sum_{t=0}^{\infty} \beta^t h_0(x, a, s) \\ &+\mathrm{E}_0 \sum_{t=0}^{\infty} \beta^t \gamma_t^j \left[h_1^j(x_t, a_t, s_t) + \sum_{n=1}^{N_j} \beta^t h_2^j(x_{t+n}, a_{t+n}, s_{t+n}) \right] \end{aligned}$$

By minimality of $\boldsymbol{\gamma}^*$, for every $\boldsymbol{\gamma} \geq 0$,

$$H_1(\mathbf{a}^*, \boldsymbol{\gamma}^* + \boldsymbol{\gamma}) \geq H_1(\mathbf{a}^*, \boldsymbol{\gamma}^*)$$

it follows that, almost surely, for every t,

$$h_1^j(x_t^*, a_t^*, s_t) + \sum_{n=1}^{N_j} \beta^t h_2^j(x_{t+n}^*, a_{t+n}^*, s_{t+n}) \geq 0$$

Identifying, as in the proof of Theorem 1, g_m^j with $h_m^j, m = 1, 2$, and noticing that, by the definition of **SPP**, $q(\mathbf{a}^*) \geq 0$, the previous inequalities show that $g(\mathbf{a}^*) \geq 0$. Now using again the minimality of $\boldsymbol{\gamma}^*$,

$$H_1(\mathbf{a}^*, \boldsymbol{\gamma}^*) \leq H_1(\mathbf{a}^*, \mathbf{0}) = 0$$

which, together with $\boldsymbol{\gamma}^* \geq 0$ and $g(\mathbf{a}^*) \geq 0$ imply that $H_1(\mathbf{a}^*, \boldsymbol{\gamma}^*) = 0$. Now suppose, there exist $\tilde{\mathbf{a}}$ satisfying $g(\tilde{\mathbf{a}}) \geq 0$ and $f(\tilde{\mathbf{a}}) > f(\mathbf{a}^*)$, then –identifying f with H_0- *it* must be that

$$H_0(\tilde{\mathbf{a}}) + H_1(\tilde{\mathbf{a}}, \boldsymbol{\gamma}^*) > H_0(\mathbf{a}^*) + H_1(\mathbf{a}^*, \boldsymbol{\gamma}^*)$$

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

4 Dynamic –saddle point– programming

Our main interest, however, is not in the existence of a solution to **SPP**, but in showing that solutions to SPP can have a recursive structure. We say that a function $\hat{W} : X \times MU \times S \to \mathcal{R}$ satisfies the *saddle point functional equation* (**SPFE**) corresponding to (**SPP**) if and only if:

$$\hat{W}(x,\, \mu,\, s) = \inf_{\gamma \geq 0} \sup_{a \in A(x,s)} \left\{ h(x,\, a,\, \mu,\, \gamma,\, s) + \beta \mathbf{E} \hat{W}(x',\, \mu',\, s') \right\}$$

$$x' = \ell(x,\, a,\, s') \text{ and } \mu' = \varphi(\mu,\, \gamma,\, s')$$

We derive the existence and uniqueness of a solution to **SPFE** adapting and extending the *contraction mapping* approach. As in maximization of dynamic problems, other approaches can be used to show the existence of a value function satisfying **SPFE**, but, as we will see, given our underlying assumptions, there is no loss of generality in using the *contraction mapping* approach. To apply it, we must first define an appropriate space of functions. We exploit the fact that the value function of **SPP** inherits from the function h the following quasi-linear structure: $W(x,\, \mu,\, s) = W_0(x, \mu,\, s) + W_1(x,\, \mu,\, s)$, with $W_0(x, \cdot,\, s)$ homogeneous of degree zero and $W_1(x, \cdot,\, s)$ homogeneous of degree one, therefore we constraint our search to functions satisfying this property. The main difficulty with the *contraction mapping* approach in our context is that the γ's and μ's are unbounded and, because of homogeneity of degree one of W_1, the corresponding value function must also be unbounded. We define a sequence of contraction mappings, parameterized by a bound K on $\|\gamma\|$. We then show that, with the assumptions of the last subsection, the **SPFE** has solutions (Proposition 4) and that the *contraction theorem* applies to our space of functions, for any K (Proposition 5). Finally, we use the fact that whenever the **SPP** has a solution the γ's are bounded to show the correspondence between **SPFE** and **SPP** (Theorem 3), and we end the Section collecting our results, which allows to relate **SPFE** and **PP** (Theorem 4).

We first define the space of "value" functions,

$$
\begin{array}{rll}
M = \{W & : & X \times \mathcal{R}_+^k \times S \to \mathcal{R} \;\; \text{s.t.} \\
& i) & W(\cdot,\cdot,\cdot) \; = \; W_0(\cdot,\cdot,\cdot) + W_1(\cdot,\cdot,\cdot) \text{ and } \forall \lambda > 0 \\
& & W_0(\cdot,\lambda\mu,\cdot) \; = \; W_0(\cdot,\mu,\cdot) \text{and } W_1(\cdot,\lambda\mu,\cdot) \; = \; \lambda W_1(\cdot,\mu,\cdot) \\
& ii) & W_j(\cdot,\,\cdot,\,s) \, \text{is continuous and bounded} \quad k = 0,1\}
\end{array}
$$

The space M is a normed vector space with the norm

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

In Appendix 2 we show that M is a nonempty and complete metric space (Lemma 1)[27].

We first fix an arbitrary positive constant K and let $K_\mu = \max\{K, K\|\mu\|\}$. We define the operator T_K on M by

$$
(T_K W)(x,\,\mu,\,s) :=: \inf_{\{\gamma \geq 0 : \|\gamma\| \leq K_\mu\}} \sup_{a \in A(x,s)} \{h(x,\,a,\,\mu,\,\gamma,\,s) + \beta E W(x',\,\mu',\,s')\}
$$
$$
\text{s.t. } x' = \ell(x,\,a,\,s') \text{ and } \mu' = \varphi(\mu,\,\gamma,\,s')
$$

If a solution to **SPP** exists (e.g., by Proposition 3) and if, as we postulate below, W corresponds to the value function of **SPP** then we can replace "infsup" by "minmax."[28]. We now study the properties of T_K (see Appendix 2). In particular, that $T_K : M \to M$ (Lemma 2) and that it satisfies Blackwell's conditions of *monotonicity* (Lemma 3) and *discounting* (Lemma 4). It is then easy to show that these conditions imply that the T_K operator satisfies the contraction property and, therefore, has a unique fixed point. More precisely,

Proposition 4 *Assume* **A1-A2** *and* **B1-B2**. *$T_K : M \to M$ is a contraction mapping.*

Proof: (See Appendix 2).

[27] See also Alvarez and Stokey (1995) for similar arguments extending the contraction mapping approach to homogeneous maps.

[28] We can also show directly (by similar arguments than the ones used in Proposition 3) that SPFE has a solution, for $W \in M$, by assuming **A1-A2** and **B1-B1b-B2** and that $W_j(\cdot,\,\cdot,\,s)$ is quasi-concave, $k = 0,1$.

The value function $W = T_K W$, defines a *policy map*[29] ψ_K such that, if $(a,\gamma) \in \psi_K(x,\, \mu,\, s)$, then for $x' = \ell(x,\, a,\, s)$ and $\mu' = \varphi(\mu,\, \gamma,\, s')$,

$$W(x,\, \mu,\, s) \;=\; h(x,\, a,\, \mu, \gamma,\, s) \,+\, \beta E W(x',\, \mu',\, s'))$$

We say that ψ_K *generates* the program $(\mathbf{a}^*, \boldsymbol{\gamma}^*)$ from the initial conditions $(x_0,\, \mu_0,\, s_0)$ if for all $n \geq 0$ and $(s_0, \ldots, s_n)$, $(a_n^*,\, \gamma_n^*) \in \psi_K\,(x_n^*,\, \mu_n^*,\, s_n)$, where $(x_0^*,\, \mu_0^*) = (x_0,\, \mu_0)$, $x_{n+1}^* = \ell(x_n^*,\, a_n^*,\, s_{n+1})$ and $\mu_{n+1}^* = \varphi(\mu_n^*,\, \gamma_n^*,\, s_{n+1})$. As usual, the fact that T_K is a contraction guarantees existence and uniqueness of a unique W that is a fixed point of this mapping, and that iterations on this mapping converge to it.

Theorem 3. Let **SPP**, satisfying **A1-A2** and **B1-B2,** have a solution. **a**) If $W : X \times \mathcal{R}_+^k \times S \to \mathcal{R}$ is the value function of **SPP,** then there exist a $\bar{K}$ such that, for all $K \geq \bar{K}$, $T_K W = W$. **b**) If $(\mathbf{a}^*, \boldsymbol{\gamma}^*)$ is generated by ψ_K from $(x_0,\, \mu_0,\, s_0)$ and, for all t, $\|\gamma_t^*\| < K_{\mu_t^*}$, then $(\mathbf{a}^*, \boldsymbol{\gamma}^*)$ solves (SPP) with initial conditions $(x_0,\, \mu_0,\, s_0)$.

Proof: We first show that $W \in M$. That it can be decomposed –as in (i)– follows from the fact that

$$\begin{aligned} & W(x_0,\, \mu_0,\, s_0) = E_0 \sum_{t=0}^{\infty} \beta^t h(x_t^*,\, a_t^*,\, \mu_t^*,\, \gamma_t^*,\, s_t) \\ = \;\; & E_0 \sum_{t=0}^{\infty} \beta^t h_0(x_t^*,\, a_t^*) + E_0 \sum_{t=0}^{\infty} \beta^t \left[\gamma_t^* h_1(x_t^*,\, a_t^*) + \mu_t^* h_2(x_t^*,\, a_t^*)\right] \\ \equiv \;\; & W_0(x_0,\, \mu_0,\, s_0) + W_1(x_0,\, \mu_0,\, s_0) \end{aligned}$$

Given this decomposition, to show that W satisfies the homogeneity properties of (i) it suffices to show that, for any $\lambda > 0$, "$(\mathbf{a}^*,\, \boldsymbol{\gamma}^*)$ is a solution to SPP with initial conditions (x, μ, s) if and only if $(\mathbf{a}^*, \lambda\, \boldsymbol{\gamma}^*)$ is a solution to SPP with initial conditions $(x, \lambda\mu, s)$." We now prove this claim.

[29]It can also be shown, using standard arguments, that the correspondence ψ is upper-hemi-continuous and, therefore, has a measurable selection (see, for example, Stokey, *et al.* (1989) *;* in particular, Theorem 7.6 p.184).

Denote by $(\mathbf{a}^*(\lambda),\, \boldsymbol{\gamma}^*(\lambda))$ a solution to SPP with initial conditions $(x, \lambda\mu, s)$, and similarly for λ'. Now,

$$\begin{aligned}
H(\mathbf{a}^*(\lambda),\, \boldsymbol{\gamma}^*(\lambda)) &= H_0(\mathbf{a}^*(\lambda),\, \boldsymbol{\gamma}^*(\lambda)) + H_1(\mathbf{a}^*(\lambda),\, \boldsymbol{\gamma}^*(\lambda)) \\
&= H_0(\mathbf{a}^*(\lambda),\, \boldsymbol{\gamma}^*(\lambda)) + \lambda\mu h_2(x, \mathbf{a}^*(\lambda)_0) \\
&\geq H_0(\mathbf{a}^*(\lambda'),\, \boldsymbol{\gamma}^*(\lambda)) + H_1(\mathbf{a}^*(\lambda'),\, \boldsymbol{\gamma}^*(\lambda)) \\
&= \frac{\lambda}{\lambda'}\left[H_0(\mathbf{a}^*(\lambda'),\, \frac{\lambda'}{\lambda}\boldsymbol{\gamma}^*(\lambda)) + H_1(\mathbf{a}^*(\lambda'), \frac{\lambda'}{\lambda}\,\boldsymbol{\gamma}^*(\lambda))\right] \\
&\quad + (1 - \frac{\lambda}{\lambda'})H_0(\mathbf{a}^*(\lambda'),\, \frac{\lambda'}{\lambda}\boldsymbol{\gamma}^*(\lambda)) \\
&\geq \frac{\lambda}{\lambda'}\left[H_0(\mathbf{a}^*(\lambda'),\, \boldsymbol{\gamma}^*(\lambda')) + H_1(\mathbf{a}^*(\lambda'),\, \boldsymbol{\gamma}^*(\lambda'))\right] \\
&\quad + (1 - \frac{\lambda}{\lambda'})H_0(\mathbf{a}^*(\lambda'),\, \boldsymbol{\gamma}^*(\lambda')) \\
&= H_0(\mathbf{a}^*(\lambda'),\, \boldsymbol{\gamma}^*(\lambda')) + \lambda\mu h_2(x, \mathbf{a}^*(\lambda')_0)
\end{aligned}$$

The first equality follows from the same argument used in the proof of Theorem 2, regarding the saddle point nature of $(\mathbf{a}^*(\lambda),\, \boldsymbol{\gamma}^*(\lambda))$ (i.e., $H_1(\mathbf{a}^*,\, \boldsymbol{\gamma}^*) = \mu h_2(x, \mathbf{a}_0^*)$); the first inequality from the maximality of $a^*(\lambda)$;the second inequality from the minimality of $\boldsymbol{\gamma}^*(\lambda')$ and the fact that, in fact, H_0 does not depend on $\boldsymbol{\gamma}^*$, and, again, the last equality follows from the saddle point nature of $(\mathbf{a}^*(\lambda'),\, \boldsymbol{\gamma}^*(\lambda'))$. Since these inequalities are satisfied for arbitrary $\lambda > 0$ and $\lambda' > 0$, it follows that, for fixed (x, μ, s),$\mathbf{a}^*(\lambda) = \mathbf{a}^*(1)$ (more precisely, that $\mathbf{a}^*(\lambda)$ is a maximal element of **SPP** with initial conditions (x, μ, s))

To see that $\boldsymbol{\gamma}^*(\lambda) = \lambda\boldsymbol{\gamma}^*(1)$ (more precisely, that $\lambda^{-1}\boldsymbol{\gamma}^*(\lambda)$ is a minimal element of SPP with initial conditions (x, μ, s)) notice that

$$\begin{aligned}
H(\mathbf{a}^*(\lambda),\, \boldsymbol{\gamma}^*(\lambda)) &\leq H(\mathbf{a}^*(\lambda), \lambda\,\boldsymbol{\gamma}^*(1)) \\
&= H_0(\mathbf{a}^*(\lambda),\, \boldsymbol{\gamma}^*(1)) + \lambda H_1(\mathbf{a}^*(\lambda),\, \boldsymbol{\gamma}^*(1)) \\
&= \lambda\left[H_0(\mathbf{a}^*(1),\, \boldsymbol{\gamma}^*(1)) + H_1(\mathbf{a}^*(1),\, \boldsymbol{\gamma}^*(1))\right] \\
&\quad + (1 - \lambda)H_0(\mathbf{a}^*(1),\, \boldsymbol{\gamma}^*(1)) \\
&\leq \lambda\left[H_0(\mathbf{a}^*(1), \lambda^{-1}\,\boldsymbol{\gamma}^*(\lambda)) + H_1(\mathbf{a}^*(1),\, \lambda^{-1}\,\boldsymbol{\gamma}^*(\lambda))\right] \\
&\quad + (1 - \lambda)H_0(\mathbf{a}^*(1),\, \lambda^{-1}\,\boldsymbol{\gamma}^*(\lambda)) \\
&= H_0(\mathbf{a}^*(1), \boldsymbol{\gamma}^*(\lambda)) + H_1(\mathbf{a}^*(1),\, \boldsymbol{\gamma}^*(\lambda)) \\
&= H(\mathbf{a}^*(\lambda),\, \boldsymbol{\gamma}^*(\lambda))
\end{aligned}$$

where, the first inequality follows from the minimality of $\boldsymbol{\gamma}^*(\lambda)$; the second inequality from the previous identity $\mathbf{a}^*(\lambda) = \mathbf{a}^*(1)$;the second inequality from the minimality of $\boldsymbol{\gamma}^*(1)$; the third equality from the homogeneity properties of H, and the last equality follows, again, from the previous identity: $\mathbf{a}^*(\lambda) = \mathbf{a}^*(1)$.

A standard generalization of the maximum principle shows the continuity of $W_k(\cdot, \cdot, s)$,while its boundedness is inherited from the h_j functions. Therefore, $W \in M$.

Now let $\bar{K}$ be the constant of Corollary to Proposition 3, then by this Corollary, for $K \geq \bar{K}$, if $(\mathbf{a}^*, \boldsymbol{\gamma}^*)$ is a solution to **SPP** with initial conditions (x, μ, s),then$\|\boldsymbol{\gamma}^*\|_\beta < K_\mu$, which implies that $\|\boldsymbol{\gamma}_0^*\| < K_\mu$. Similarly, at (x_t^*, μ_t^*, s_t),the bound on SPP implies that $\|\boldsymbol{\gamma}_t^*\| < K_{\mu_t^*}$. It follows that, for $K \geq \bar{K}$, the bound on T_K is not binding for **SPP** solutions. Uniqueness of the contraction map guarantees that $T_K W = W$.

To see the second part of the theorem, let $(\mathbf{a}^*, \boldsymbol{\gamma}^*)$ be generated by $\psi_{K'}$ from (x_0, μ_0, s_0),then for some $\widehat{W} \in M$, $\widehat{W}(x_0, \mu_0, s_0) = T_K \widehat{W}(x_0, \mu_0, s_0)$ for all $K \geq K'$ since, by assumption, for all t, $\|\gamma_t^*\| < K_{\mu_t^*}$. But by (**a**), for $K \geq \overline{K}$, $T_K W = W$, where,W is the value function of SPP. By uniqueness, it follows that $W(x_0, \mu_0, s_0) = \widehat{W}(x_0, \mu_0, s_0) = E_0 \sum_{t=0}^{\infty} \beta^t h(x_t^*, a_t^*, \mu_t^*, \gamma_t^*, s_t)$. That is, $(\mathbf{a}^*, \boldsymbol{\gamma}^*)$ is a solution to **SPP** with initial conditions (x_0, μ_0, s_0).■

4.1 A final theorem

We are finally in a position to recast our results in a comprehensive theorem, which follows from the previous results, in particular, Theorems 1-3, and relates **SPFE** and **PP**.

Theorem 4. a) Let **PP** satisfy **A1**-**A3**, **A4** & **A5** and derive the corresponding **SPP**. If there exist a K and a $(\mathbf{a}^*, \boldsymbol{\gamma}^*)$ generated (from initial conditions $(x_0, 0, s_0)$) by a policy ψ_K corresponding to a value function W satisfying **SPFE**, with the property that, for all t, $\|\gamma_t^*\| < K_{\mu_t^*}$, then $\mathbf{a}^*$ solves **PP** with initial conditions (x_0, s_0), and value $V(x_0, s_0)$, Furthermore, $W\ (x_0,\ 0,\ s_0) = V(x_0, s_0)$.

b) Assume **A1**-**A2** and **A3**, **A4**, **A4b** & **A5**. **PP** has a solution, and if $\mathbf{a}^*$ solves **PP** with initial conditions (x_0, s_0), and value $V(x_0, s_0)$, then there exist a K and a $(\mathbf{a}^*, \boldsymbol{\gamma}^*)$ generated by a policy ψ_K (with initial conditions $(x_0, 0, s_0)$) derived from a value function W satisfying **SPFE**, with the property that, for all t, $\|\gamma_t^*\| < K_{\mu_t^*}$. Furthermore, W $(x_0,\ 0,\ s_0) = V(x_0, s_0)$.

Proof: **a)** By construction, the corresponding SPP satisfies A1-A2 and B1-B2. Let $(\mathbf{a}^*, \boldsymbol{\gamma}^*)$ be generated from ψ_K , with the bounds $K_{\mu_t^*}$ never binding, then by Theorem 3, $(\mathbf{a}^*, \boldsymbol{\gamma}^*)$ solves SPP, with initial conditions $(x_0, 0, s_0)$ and, by Theorem 2 , $\mathbf{a}^*$ is a solution to PP satisfying W $(x_0,\ 0,\ s_0) = V(x_0, s_0)$.

b) That **PP** has a solution follows from Proposition 1. If $\mathbf{a}^*$ solves PP with initial conditions (x_0, s_0) and we assume A1-A4, A4b & A5, by Theorem 1 there exist a $\boldsymbol{\gamma}^*$ such that $(\mathbf{a}^*, \boldsymbol{\gamma}^*)$ is a solution to a SPP (derived from PP) with initial conditions $(x_0, 0, s_0)$. By Theorem 3 the value function of SPP, W, satisfies SPFE for $K \geq \bar{K}$ and $T_K W = W$. Therefore, there exist a policy map ψ_K generating $(\mathbf{a}^*, \boldsymbol{\gamma}^*)$. ■

In summary, as we have discussed, many economic problems take the PP form. We have shown how to transform them in the SPP form and we have provided conditions guaranteeing the existence of solutions to SPP, and -weaker- conditions, under which, a solution to SPP is also a solution to PP. More interestingly from a computational point of view, we have also shown that, under our assumptions, all SPP solutions satisfy a saddle point functional equation (SPFE). That is, in applications of our theory one only has to check that PP satisfies our conditions (i.e., A1-A3, A4 & A5) to be able to guarantee by Theorem 4a) that the solution of SPFE is a solution to the original PP problem. In fact, we provide conditions guaranteeing that any SPP solution can be achieved as a solution to SPFE. With global convexity assumptions (i.e., assuming A4b too), we can be sure that *all* **PP** *solutions* can be found by solving SPFE, without such convexity assumptions it may be that there are solutions to PP that are not solutions to the corresponding SPP formulation. Nevertheless, provided that SPP has a solution, then the maximal value $V(x_0, s_0)$ is achieved by our recursive characterization in terms of a SPFE..

5 Conclusions and extensions

We have shown that a large class of problems with implementability constraints can be analyzed by an equivalent recursive saddle point problem. This saddle point problem obeys a saddle point functional equation, which is a version of the Bellman equation. This approach works for a very large class of models with incentive constraints, limits in the budget constraint, optimal policy, optimal regulation, etc. This means that a unified framework can be provided to analyze all these models. Instead of having to write optimal contracts as history-dependent contracts one can write them as a stationary function of few state (and co-state) variables. This means, for example, that the time-inconsistency problem does not complicate considerably the numerical solution to this problem, only a few co-state variables need to be added, and computation of the solution is greatly simplified.

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

APPENDIX 1 (Proofs of Section 3)

Proof of Proposition 2 The existence of a functional $\gamma_b^* \in ba_+$ that satisfies

$$L(\mathbf{a}^*, \boldsymbol{\gamma}_b) :\geq: L(\mathbf{a}^*, \boldsymbol{\gamma}_b^*) :\geq: L(\mathbf{a}, \boldsymbol{\gamma}_b^*) \tag{15}$$

for all $\mathbf{a} \in \mathcal{A}$ and $\boldsymbol{\gamma}_b \in ba_+$ follows from the standard theory of constrained optimization in linear vector spaces (see, for example, (Luenberger (1969), Section 8.3, Theorem 1 and Corollary 1)). To obtain this result we make use of the following facts: *i)* a solution to **PP** exists (Proposition 1); *ii)* as we already mentioned, by assumptions **A4** and **A5** $g(\cdot)$ maps into $\mathcal{L}_\infty$ and $\mathcal{L}_{\infty,+}$ has a non empty interior; *iii)* by assumption **A3**, $f(\cdot)$ is continuous and quasiconcave, and *iv)* by assumption **A5**, a Slater interiority condition is satisfied.

We have to show that inequalities (15) are also satisfied with multipliers that are countably additive (i.e., in $\mathcal{L}_{1,+}$).

Given an initial condition (x_0, s_0), we can model the exogenous uncertainty as an infinite branching process from s_0. Abusing notation, let (for the remaining of the proof) $S_0 = \{s_0\}$ and S_t be the set of possible values of s_t following s_0. If there are $n = k + 1$ constraints in period t, let $Z = \cup_0^\infty (S_t \times n)$. Since, by assumption **A1**, $\{s_t\}$ is a Markovian process, there is a well defined measure space $(Z, \mathcal{Z}, \nu)$. That is, $g : \mathcal{A} \rightarrow \mathcal{L}_\infty(Z, \mathcal{Z}, \nu)$, and $\boldsymbol{\gamma}_b \in ba_+(Z, \mathcal{Z}, \nu)$. Let $\gamma_{b,t}(A) = \boldsymbol{\gamma}(A \cap (S_t \times n))$.

Second, we recall some mathematical facts (already used in Bewley (1972)). By Yosida-Hewitt decomposition, given $\boldsymbol{\gamma} \in ba_+$, there exist unique $\gamma_{c,t} \geq 0$ and $\gamma_{p,t} \geq 0$ such that $\gamma_{c,t}$ is countable additive and $\gamma_{p,t}$ is purely finitely, satisfying: $\gamma_{b,t} = \gamma_{c,t} + \gamma_{p,t}$. Furthermore, for every $\epsilon_t > 0$, there exist $A_t \in \mathcal{Z}$, such that $\gamma_{c,t}(A_t) < \epsilon_t$ and $\gamma_{p,t}(Z \setminus A_t) = 0$. It follows that, if $\{s : a_t^{(n)}(s) \neq a_t(s)\} \subseteq Z \setminus A_t^{(n)}$, and $\sum_t \epsilon_t^{(n)} \rightarrow 0$ (as $n \rightarrow \infty$; which can be achieved by an appropriate choice of $\{\epsilon_t^{(n)}\}$), then $\lim_n \boldsymbol{\gamma}_c(\{s : \mathbf{a}^{(n)}(s) \neq \mathbf{a}(s)\} = 0$ and, if $f(\cdot)$ is continuous in probability (i.e., in the P-topology), then $f(\mathbf{a}^{(n)}) \rightarrow f(\mathbf{a})$.

We now use these facts, and the interiority assumption **A5**, to show that $\boldsymbol{\gamma}_c^*$ is, in fact, a supporting Lagrange multiplier.

Consider first the left inequality of (15). Since $g(\mathbf{a}^*) \geq 0$, it follows that $\boldsymbol{\gamma}_b^* g(\mathbf{a}^*) = 0$. In the Yosida-Hewitt decomposition both terms are nonnegative, therefore $\boldsymbol{\gamma}_c^* g(\mathbf{a}^*) = 0$. On the other hand, for all $\boldsymbol{\gamma} \in \mathcal{L}_1+$, $\boldsymbol{\gamma} g(\mathbf{a}^*) \geq 0$,. These last two facts show the left inequality of the saddle point condition (14).

Now we show the right inequality of (14). Suppose for some $\mathbf{a} \in \mathcal{A}$, $f(\mathbf{a}) + \boldsymbol{\gamma}_c^* g(\mathbf{a}) > f(\mathbf{a}^*) + \boldsymbol{\gamma}_c^* g(\mathbf{a}^*)$. For $t > 0$, let $F_t^{(n)} = \cup_{r=1}^t A_r^{(n)}$, where the sets $A_r^{(n)}$ are ordered –according to the branching stochastic process starting from s_0– and satisfy the above Hewitt-Yosida decomposition conditions. Let $F_0^{(n)} = \emptyset$. Then, a new contract $\mathbf{a}^{(n)}$ can be defined by $a_0^{(n)} = a_0$, and, for $t > 0$,

$$a_t^{(n)}(s) = \begin{cases} a_t(s) & \text{if } s \notin F_t^{(n)} \\ \hat{a}_t(s) & \text{if } s \in F_t^{(n)} \setminus F_{t-1}^{(n)} \end{cases}$$

where $\{\hat{a}_{t+r}(s)\}_{r=0}^{\infty}$ is the interior program –of assumption **A5**– starting from $(x_t(s), s_t)$. Therefore, we have that,

$$\begin{aligned} f(\mathbf{a}^{(n)}) + \boldsymbol{\gamma}_b^* g(\mathbf{a}^{(n)}) &\geq f(\mathbf{a}^{(n)}) + \boldsymbol{\gamma}_c^* g(\mathbf{a}^{(n)}) \\ &\to f(\mathbf{a}) + \boldsymbol{\gamma}_c^* g(\mathbf{a}) \end{aligned}$$

The inequality follows from the fact that $\boldsymbol{\gamma}_p \geq 0$ and on sets with purely finitely positive measure the program is, by construction, interior. The convergence property follows from the P-continuity of f, the boundedness assumption **A4** and the construction of $\mathbf{a}^{(n)}$. It follows that, for n large enough,

$$\begin{aligned} f(\mathbf{a}^{(n)}) + \boldsymbol{\gamma}_b^* g(\mathbf{a}^{(n)}) &> f(\mathbf{a}^*) + \boldsymbol{\gamma}_c^* g(\mathbf{a}^*) \\ &= f(\mathbf{a}^*) + \boldsymbol{\gamma}_b^* g(\mathbf{a}^*) \end{aligned}$$

which contradicts (15) and proves that the saddle point condition (14) is satisfied (with $\boldsymbol{\gamma}_c^*$).

Finally, to see the last statement of the proposition, notice that

$$L(\mathbf{a}^*, \boldsymbol{\gamma}_c^*) = f(\mathbf{a}^*) = V(x_0, s_0),$$

where the first equality follows from the fact, already proved, that $\boldsymbol{\gamma}_b^* g(\mathbf{x}^*) = \boldsymbol{\gamma}_c^* g(\mathbf{x}^*) = 0$ and the second equality follows by definition∎

Proof of Proposition 3: Our goal is to show the existence of a solution to the following problem:

SPP

$$W = \inf_{\boldsymbol{\gamma}\geq 0} \sup_{q(\mathbf{a})\geq 0} H(\mathbf{a}, \boldsymbol{\gamma})$$

We fix the initial conditions (x_0, μ_0, s_0) and proceed in two steps. First we show that if we further restrict the set of feasible $\boldsymbol{\gamma}$ sequences, the corresponding $\mathbf{SPP}_m$ problem has a solution, and second we show that, with our interiority assumptions, such a restriction can be made without loss of generality. The truncated problem is

$\mathbf{SPP}_m$

$$\inf_{\{\boldsymbol{\gamma}:\boldsymbol{\gamma}\geq 0 \text{ and } \|\boldsymbol{\gamma}\|_\beta \leq m\}} \sup_{q(\mathbf{a})\geq 0} H(\mathbf{a}, \boldsymbol{\gamma})$$

where $\|\boldsymbol{\gamma}\|_\beta = \sum_{t=0}^{\infty} \beta^t \|\gamma_t\|$.

Before we proceed, we collect a couple of facts that we will use in the proof. *i)* if $\mathbf{a}^n \to \mathbf{a}$, in probability, then, given an initial condition (x_0, s_0), $\mathbf{x}^n \to \mathbf{x}$ in probability, where, $x_0^n = x_0$and, for $t \geq 0$, $x_{t+1}^n = \ell(x_t^n, a_t^n, s_{t+1})$,which in turn implies that $\{h_j(x_t^n, a_t^n)\} \to \{h_j(x_t, a_t)\}$, $j = 0, 1, 2$ in the P-topology; *ii)* by (the second part of) **B2** if $\|\gamma\| \leq m$ $\|\varphi(\mu, \gamma, s)\| \leq m+ \|\mu\|$.

We now decompose the problem as

$$\begin{aligned} R^1(\boldsymbol{\gamma}) &= \max\nolimits_{q(\mathbf{a})\geq 0} H(\mathbf{a}, \boldsymbol{\gamma}) \\ &\text{and} \\ R_m^2(\mathbf{a}) &= \min\nolimits_{\{\boldsymbol{\gamma}:\boldsymbol{\gamma}\geq 0 \text{ and } \|\boldsymbol{\gamma}\|\leq m\}} H(\mathbf{a}, \boldsymbol{\gamma}) \end{aligned}$$

We first consider the existence of maximal elements $\mathbf{a}^* \in R^1(\boldsymbol{\gamma}) \subset \mathcal{L}_\infty$ By assumption **B2** the set of feasible solutions is nonempty. By assumptions **A1** and **A2,** and following the same argument that in the proof of Proposition 1, $\{\mathbf{a} : q(\mathbf{a}) \geq 0\}$ is compact in the P-topology. By fact *(i)* $H(\cdot, \boldsymbol{\gamma})$ is P-continuous, whenever $\{(\beta^t, \beta^t\gamma_t, \beta^t\mu_t)\} \in \mathcal{L}_1$, which, by fact *(i),* it is the case if $\|\boldsymbol{\gamma}\| \leq m$; therefore, $R^1(\boldsymbol{\gamma})$ is non-empty. Now, consider the existence of minimal elements $\boldsymbol{\gamma}^* \in R_m^2(\mathbf{a})$. Given that $\beta \in (0, 1)$,the set $\left\{(\beta^t, \beta^t\gamma_t, \beta^t\mu_t) : \|\gamma\|_\beta \leq m \text{ and } \|\mu\|_\beta \leq M\right\}$ is norm bounded, pointwise closed and convex. i.e., it is $\sigma(\mathcal{L}_1, \mathcal{L}_\infty)$ compact. Furthermore, given $(\mathbf{a}, \mathbf{x}) \in \mathcal{L}_\infty$, for $j = 0, 1, 2$, $\{h_j(x_t, a_t)\} \in \mathcal{L}_\infty$. It follows that $H(\mathbf{a}, \cdot)$ is $(\sigma(\mathcal{L}_1, \mathcal{L}_\infty))$ continuous in $\boldsymbol{\gamma}$, and that $R^2_m(\boldsymbol{\gamma})$ is non-empty. Compactness of the constraint sets and the continuity and quasiconcavity properties of $H(\cdot, \cdot)$ (for fixed (x_0, μ_0, s_0)), imply that $R_m(\cdot, \cdot) \equiv (R^1(\cdot), R^2_m(\cdot))$ is a convex-valued upper-hemi continuous correspondence (i.e, has a closed graph) mapping a convex, compact set in itself. It follows from a fixed point theorem that there exist $(\mathbf{a}^*, \boldsymbol{\gamma}^*) \in R_m(\mathbf{a}^*, \boldsymbol{\gamma}^*)$.

Now, we show that for large m, $(\mathbf{a}^*, \boldsymbol{\gamma}^*)$ is also a fixed point of the untruncated problem.

Let $\hat{\mathbf{a}}$ (and the corresponding $\hat{\mathbf{x}}$) be the interior program of assumption **B3**, then notice that

$$\begin{aligned} H(\hat{\mathbf{a}}, \boldsymbol{\gamma}^*) &= E_0 \sum_{t=0}^{\infty} \beta^t h_0(\hat{x}_t, \hat{a}_t) + \mu_0 h_2(x_0, \hat{a}_0) \\ &\quad + E_0 \sum_{t=0}^{\infty} \beta^t \left[\gamma^*_t h_1(\hat{x}_t, \hat{a}_t) + \beta \mu^*_{t+1} h_2(\hat{x}_{t+1}, \hat{a}_{t+1})\right] \\ &\geq E_0 \sum_{t=0}^{\infty} \beta^t h_0(\hat{x}_t, \hat{a}_t) + \mu_0 h_2(x_0, \hat{a}_0) + \epsilon \, \|\boldsymbol{\gamma}^*\|_\beta \end{aligned}$$

Therefore,

$$\begin{aligned} 0 &\leq H(\mathbf{a}^*, \boldsymbol{\gamma}^*) - H(\hat{\mathbf{a}}, \boldsymbol{\gamma}^*) \leq H(\mathbf{a}^*, \boldsymbol{\gamma}_{\mu_0}) - H(\hat{\mathbf{a}}, \boldsymbol{\gamma}^*) \\ &\leq E_0 \sum_{t=0}^{\infty} \beta^t \left[h_0(x^*_t, a^*_t) - h_0(\hat{x}_t, \hat{a}_t)\right] + \mu_0 \left[h_2(x_0, a^*_0) - h_2(x_0, \hat{a}_0)\right] \\ &\quad + E_0 \sum_{t=0}^{\infty} \beta^t \left[\gamma_{\mu_0} h_1(x^*_t, a^*_t) + \beta \mu_0 h_2(x^*_{t+1}, a^*_{t+1})\right] - \epsilon ||\boldsymbol{\gamma}^*||_\beta \\ &\leq B \cdot \max\{1, \|\mu_0\|\} - \epsilon ||\boldsymbol{\gamma}^*||_\beta \end{aligned}$$

The first inequality follows from the maximality of $\mathbf{a}^*$; the second from minimality of $\boldsymbol{\gamma}^*$;the third from the previous inequalities, and from the boundedness assumption **B1**. Therefore, $||\boldsymbol{\gamma}^*||_\beta \leq \max\{1, \|\mu_0\|\} B/\epsilon$.

Now let $m \geq \max\{1, \|\mu_0\|\} \bar{K} \equiv \max\{1, \|\mu_0\|\} 2B/\epsilon$, and $(\mathbf{a}^*, \boldsymbol{\gamma}^*) \in R_m(\mathbf{a}^*, \boldsymbol{\gamma}^*)$. Suppose there exist a $\tilde{\boldsymbol{\gamma}} \geq 0$, $||\tilde{\boldsymbol{\gamma}}||_\beta > m$, satisfying $H(\mathbf{a}^*, \boldsymbol{\gamma}^*) > H(\mathbf{a}^*, \tilde{\boldsymbol{\gamma}})$. Let $\bar{\boldsymbol{\gamma}} \geq 0$ be such that, for all $t \geq 0$ $\bar{\gamma}_t = \alpha\gamma_t^* + (1-\alpha)\tilde{\gamma}_t$, for some $\alpha \in (0, 1)$, and $||\bar{\boldsymbol{\gamma}}||_\beta \leq m$. But then, by **B2**

$$\begin{aligned} & H(\mathbf{a}^*, \boldsymbol{\gamma}^*) - H(\mathbf{a}^*, \bar{\boldsymbol{\gamma}}) \\ = \; & (1-\alpha)[H(\mathbf{a}^*, \boldsymbol{\gamma}^*) - H(\mathbf{a}^*, \bar{\boldsymbol{\gamma}})] \\ > \; & 0 \end{aligned}$$

and this contradicts the fact that $\boldsymbol{\gamma}^* \in R_m^2(\mathbf{a}^*)$.

Finally, to see the uniqueness of the value. Let $(\mathbf{a}^*, \boldsymbol{\gamma}^*)$ and $(\mathbf{a}', \boldsymbol{\gamma}')$ be two solutions to **SPP**, then by applying the saddle point property to both solutions we obtain

$$H(\mathbf{a}', \boldsymbol{\gamma}') \geq H(\mathbf{a}^*, \boldsymbol{\gamma}') \geq H(\mathbf{a}^*, \boldsymbol{\gamma}^*) \geq H(\mathbf{a}', \boldsymbol{\gamma}^*) \geq H(\mathbf{a}', \boldsymbol{\gamma}')$$

■

Proof of Corollary to Proposition 3: It follows from the proof to Proposition 3 that, if assumptions there exist a $\bar{K}$ such that the solution $(\mathbf{a}^*, \boldsymbol{\gamma}^*)$ to **SPP** satisfies $\|\boldsymbol{\gamma}^*\|_\beta < \bar{K} \cdot \max\{1, \|\mu_0\|\}$ ■

APPENDIX 2 (Proofs of Section 4)

Lemma 1. *M* is a nonempty complete metric space.

Proof: That it is non-empty is trivial. That every Cauchy sequence $\{W^n\} \in M$ converges to a function $\tilde{W}$ satisfying *ii*) follows from standard arguments (see, for example, Stokey, *et al.* (1989), Theorem 3.1 and Lemma 9.5); these arguments apply to both components W_j^n, $j = 0, 1$. To see that the homogeneity properties are also satisfied, for any (x, μ, s) and $\lambda > 0$, let $\lambda^0 = 1$ and $\lambda 1 = \lambda$, then, for $j = 0, 1$,

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

■

Lemma 2. The operator T_K maps M into itself.

Proof:

$$(T_K W)(x,\, \mu,\, s) = \left[h_0(x,\, a^*,\, s) \,+\, \beta E W_0(x^{*'},\, \mu^{*'},\, s')\right]$$
$$+ \left[\gamma^* h_1(x,\, a^*,\, s) \,+\, \mu h_2(x,\, a^*,\, s) \,+\, \beta E W_1(x^{*'},\, \mu^{*'},\, s')\right]$$

therefore,

$$\|(T_K W)(x,\, \mu,\, s)\| \leq \|h_0(x,\, a^*,\, s)\,\| + \beta \left\| W_0(x^{*'},\, \mu^{*'},\, s') \right\|$$
$$+ \max\{1, \|\,\mu\|\}\, K\, \|h_1(x,\, a^*,\, s)\,\| + \|\,\mu\|\, \|h_2(x,\, a^*,\, s)\|$$
$$+ \beta(\max\{1, \|\,\mu\|\}\, K + \|\,\mu\|) \left\| W_1(x^{*'},\, \frac{\mu^{*'}}{\|\,\mu^{*'}\|},\, s') \right\|$$

It follows that the boundedness condition of $ii)$ is satisfied. A routine generalization of *the maximum principle* (see, for example, Stokey, *et al.* (1989)) to this saddle point case, shows that $(TW)(\cdot,\, \cdot,\, s)$ is continuous. To see that the homogeneity properties are satisfied, consider $(x,\, \lambda\mu,\, s)$, with $\lambda > 0$, and a corresponding solution $(a^*_\lambda,\, \gamma^*_\lambda)$. Let $\gamma^*_1 = \lambda^{-1}\gamma^*_\lambda$, then

$$\begin{aligned}
&(TW)(x,\, \lambda\mu,\, s) \,=\, (TW)_0(x,\, \lambda\mu,\, s) \,+\, (TW)_1(x,\, \lambda\mu,\, s) \\
&= \left[h_0(x,\, a^*_\lambda,\, s) \,+\, \beta E W^1(x^{*'}_\lambda,\, \varphi(\mu, \gamma^*_\lambda),\, s')\right] \\
&+ \;\left[\gamma^*_\lambda h_1(x,\, a^*_\lambda,\, s) \,+\, \lambda\mu h_2(x,\, a^*_\lambda,\, s) \,+\, \beta E W_1(x^{*'}_\lambda,\; \varphi(\lambda\mu, \gamma^*_\lambda),\, s')\right] \\
&= \left[h_0(x,\, a^*_\lambda,\, s) \,+\, \beta E W_0(x^{*'}_\lambda,\, \varphi(\mu, \gamma^*_1),\, s')\right] \\
&\;+\lambda\left[\gamma^*_1 h_0(x,\, a^*_\lambda,\, s) \,+\, \mu h_2(x,\, a^*_\lambda,\, s) \,+\, \beta E W_1(x^{*'}_\lambda,\, \varphi(\mu, \gamma^*_1),\, s')\right] \\
&= (TW)_0(x,\, \mu,\, s) \,+\, \lambda(TW)_1(x,\, \mu,\, s)
\end{aligned}$$

■

Lemma 3 (monotonicity) Let $F, G \in$ M be such that $F \leq G$, then $(T_K F) \leq (T_K G)$.

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

$$\max_{a \in A(x,s)} \{h(x, a, \mu, \gamma,\, s) + \beta E F(\ell(x,\, a,\, s), \mu',\, s')\}$$
$$\leq \max_{a \in A(x,s)} \{h(x, a, \mu, \gamma,\, s) + \beta E G(\ell(x,\, a,\, s), \mu',\, s')\}$$

It follows that

$$\min_{\{\gamma\geq 0:||\gamma||\leq K_\mu\}} \max_{a\in A(x,s)} \{h(x,a,\mu,\gamma,\, s)+\beta EF(\ell(x,\, a,\, s),\varphi(\mu,\, \gamma,\, s),\, s')\}$$
$$\leq \min_{\{\gamma\geq 0:||\gamma||\leq K_\mu\}} \max_{a\in A(x,s)} \{h(x,a,\mu,\gamma,\, s)+\beta EG(\ell(x,\, a,\, s),\varphi(\mu,\, \gamma,\, s),\, s')\}$$

■

In our context, if $F \in M$ and $a \in \mathcal{R}$, we define the function $F + a \in M$ by $(F + a)(x,\, \mu,\, s) = F(x,\, \mu,\, s) + a$.

Lemma 4 (discounting) For all $W \in$ M , and $a \in \mathcal{R}_+$, $T_K(W + a) \leq T_K W + \beta a$.

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

$$\max_{a\in A(x,s)} \{h(x,a,\mu,\gamma,\, s)+\beta E(W+a)(\ell(x,\, a,\, s),\varphi(\mu,\, \gamma,\, s),\, s')\}$$
$$= \max_{a\in A(x,s)} \{h(x,a,\mu,\gamma,\, s)+\beta EW(\ell(x,\, a,\, s),\varphi(\mu,\, \gamma,\, s),\, s')+\beta a\}$$
$$= \max_{a\in A(x,s)} \{h(x,a,\mu,\gamma,\, s)+\beta EW(\ell(x,\, a,\, s),\varphi(\mu,\, \gamma,\, s),\, s')\}+\beta a$$

Now, using these equalities and the above definition for $F + a$,

$$T_K(W\ +\ a)(x,\mu,\, s)$$
$$= \min_{\{\gamma\geq 0:||\gamma||\leq K_\mu\}} \max_{a\in A(x,s)} \{h(x,a,\mu,\gamma,\, s)+\beta E(W+a)(\ell(x,\, a,\, s),\varphi(\mu,\, \gamma,\, s),\, s')\}$$
$$= \min_{\{\gamma\geq 0:||\gamma||\leq K_\mu\}} \max_{a\in A(x,s)} \{h(x,a,\mu,\gamma,\, s)+\beta EW(\ell(x,\, a,\, s),\varphi(\mu,\, \gamma,\, s),\, s')\}+\beta a$$
$$= (T_K W+\beta a)(x,\mu,\, s)$$

We have shown that $T_K(W\ +\ a) \leq T_K W\ +\ \beta a$■

Proof of Proposition 4: The argument is standard. We show that the contraction property is satisfied. Let $F, G \in M$, then, using the homogeneity property of the functions in M, for any $(x,\, \mu,\, s)$,

$$\begin{aligned} F(x,\, \mu,\, s) &= G(x,\, \mu,\, s)+[F(x,\, \mu,\, s)-G(x,\, \mu,\, s)] \\ &\leq G(x,\, \mu,\, s)+|F(x,\, \mu,\, s)-G(x,\, \mu,\, s)| \end{aligned}$$

That is, $F \leq G + ||F - G||$. By the *monotonicity* and the *discounting* properties, it follows that $T_K F \leq T_K G + \beta ||F - G||$. But now, reversing the roles of F and G we obtain that

$$||T_K F - T_K G|| \leq \beta ||F - G||$$

Since $0 < \beta < 1$ we have that T_K is a contraction mapping.

APPENDIX 3 (A1-A5 in examples 1 and 2)

Example 1

It is easy to check that, with more structure on the model, assumptions A1–A4, A4b & A5 are all satisfied. For A1 we only need to assume that the support of (θ_t, ω_t) is contained in a bounded set for all t. The Feller property is satisfied, for example, under the usual cases that $\{\theta_t, \omega_t\}_{t=0}^{\infty}$ takes only discrete values or the conditional distribution of (θ_t, ω_t) conditional on $(\theta_{t-1}, \omega_{t-1})$ is continuous with respect to the latter. For the next assumptions, if $0 < \delta < 1$, $F : \mathcal{R}_+ \times \mathcal{R} \to \mathcal{R}_+$ and $\lim_{\overline{k} \to \infty} F_k(\overline{k}, \theta) < 1 - \delta$ almost surely in θ (where F_k is the derivative with respect to k) we know that there exists a $\overline{k} < \infty$ such that $k_t < \overline{k}$; also, if F and u are assumed to be continuous in $[0,\overline{k}]$, then the mappings A, ℓ and r are clearly continuous and bounded in this interval, so A2 and A3 are satisfied. The assumption that u is continuous plus the assumptions on the shock s we have introduced above guarantee that the value function of autarky v_j^a is continuous and bounded for both agents, which guarantees A4. Furthermore, if u and F are assumed quasi-concave, then assumption A4b is also satisfied.[30]

Finally, for the interiority condition A5, we first assume that $k_0 \geq k^l$ for $k^l > 0$ sufficiently small[31], and that for all $k \geq k^l$ and each possible value of the technology shock $F(k, \theta) \geq c^l + \delta k^l$ for some $c^l > 0$. This is satisfied, for example, if F is increasing, the usual Inada condition $F'(0) = \infty$ holds.

[30] Proving **A4b** is very easy in this example because of the fact that v^a does not depend on endogenous variables. In models where the value of autarky depends on the capital, such as the model of Marcet and Marimon (1992), proving quasi-concavity is more complicated.

[31] Or, alternatively, by introducing the constraint $k_t \geq k^l$ in the technology A.

Under such conditions, a feasible sequence for consumptions and investments can be constructed to satisfy

$$\widehat{i}_t = \delta k^l, \qquad \widehat{k}_t = (1-\delta)\widehat{k}_{t-1} + \widehat{i}_t \qquad \widehat{c}_{jt} = \omega_{jt} + \frac{F(\widehat{k}_t, \theta) - \delta k^l}{3} \tag{16}$$

for all t. This sequence satisfies $\widehat{k}_t \geq k^l$ and, therefore, $\widehat{c}_{jt} \geq \omega_{jt} + c^l/3$. Since $c^l/3$ units of production are thrown away every period, we have $\widehat{c}_{1t} + \widehat{c}_{2t} + \widehat{i}_t \leq F(\widehat{k}_t, \theta_t) + \omega_{1t} + \omega_{2t} + c^l/3$ and the choice variables are in the interior of the feasible set. Finally, define the function $\zeta(\overline{c}) = \min_{\omega \in S^{\omega_j}} u(\omega + \overline{c}) - u(\omega)$, where S^{ω_j} is the support of all ω_{jt} and assume that, for any $\overline{c} > 0$, we have $\zeta(\overline{c}) > 0$.[32] Then we have that $u(\widehat{c}_{jt}) - u(\omega_{jt}) \geq \zeta(c^l/3)$, and assumption A5 is satisfied for

$$\varepsilon = \frac{1}{1-\beta}\zeta(c^l) \tag{17}$$

Therefore, we conclude that under mild assumptions on the technology, preferences and endowments A1-A4, A4b & A5 are satisfied. Theorem 4 guarantees that all solutions to the planner's problem (PP) can be found by solving SPFE and viceversa. If we relax the quasi-concavity assumption (for example, if F has an interval of increasing returns) Theorem 4a still guarantees that solutions to the SPFE are solutions to the PP.

Example 2

Assumptions A1–A4 can be dealt with in a similar manner as Example 1, so we will not repeat them here.

To guarantee assumption A4b and A5, however, is not easy in this example. The first difficulty in is that, because of the 'equality' sign in equation (11) the set of allocations satisfying this equation is not convex, and the interiority assumption can not be satisfied. We will proceed by replacing the equality in the PP by a weak inequality; if we can then show that in the optimum the planner chooses an allocation where the equation satisfied as

[32] That $\zeta(\overline{c}) > 0$ can be guaranteed, for example, if the derivative of u is bounded away from zero in the support of the ω's. Most applications satisfy this property.

an equality, we can be sure that the optimum is the same as with (11) and that we are solving the model of interest[33].

Now we have to decide if we write the inequality as $\geq$ or as $\leq$. Consider the case $\geq$ and let the full optimum without distortionary taxes be denoted by $\{\widetilde{c}_t, \widetilde{k}_t\}_{t=0}^{\infty}$. Clearly,

$$u'(\widetilde{c}_t) = \beta E_t \left[u'(\widetilde{c}_{t+1})(\ \widetilde{r}_{t+1} + 1 - \delta)\right] \geq \beta E_t \left[u'(\widetilde{c}_{t+1}) \left(\ \widetilde{r}_{t+1}(1 - \widetilde{\tau}_{t+1}) + 1 - \delta\right)\right],$$

where the first equality is a property of the full-optimum and the inequality follows from the fact that tax rates are positive. Hence, changing the equality in (11) to a $\geq$ would make the first best feasible, the solution would be the first optimum, which is not the same as the Ramsey equilibrium. So, this option does not deliver a solution equivalent to the solution under (11).

Now, let us consider the case of replacing the equality with a $\leq$ to consider the restriction

$$u'(c_t) \leq \beta E_t \left[u'(c_{t+1}) \left(\ r_{t+1}(1 - \tau_{t+1}) + 1 - \delta\right)\right]. \tag{18}$$

This would be the first order condition if the consumer faced a constraint

$$k_t \leq k_t^U, \tag{19}$$

where k_t^U is an upper bound on capital imposed on the consumer. The Euler equation (18), therefore, corresponds to a policy environment where the government has the ability to impose some upper bounds on capital accumulation k_t^U on the consumer, and the policy instruments available to the planner are now $\{k_t^U, \tau_t\}$. Any sequence $\{c_t, k_t, \tau_t, c_t^g\}$ that satisfies (18) can be implemented by a government policy that sets $k_t^U = k_t$ in periods and realizations when the inequality is satisfied as strict inequality and k_t^U very large if (18) is satisfied as equality. It is clear that the planner will choose sequences where (18) is satisfied as an equality, since the equilibrium with distorting taxes has underaccumulation of capital relative to the full optimum $\left\{\widetilde{k}_t\right\}_{t=0}^{\infty}$. This implies that the planner facing restrictions (18) and

[33]Note that a similar approach is used in standard general equilibrium theory, where feasibility constraints written as an equality usually do not define a convex set. Often, the equality is replaced with a weak inequality and the appeal to non-satiated preferences guarantees that the feasibility constraint is also satisfied as equality.

(19) will not choose a sequence where k_t^U is binding, and the government will act so that (18) is satisfied as an equality. Then, the optimum under (18) is equivalent to the Ramsey equilibrium.

Under this modification, it is clear that the interiority condition A5 is satisfied, since the planner could choose a set of upper bounds to the capital stocks that are binding. This shows that, in this example, the sufficiency of SPFE is guaranteed.

However, in order to be sure that SPFE has a solution we have to make sure that the solution to PP is also a SPP solution. As we have seen this is guaranteed if our convexity assumptions (**A4b** or **B1b**) are satisfied[34]. Nevertheless, virtually all studies of Ramsey equilibria proceed by analyzing the lagrangean directly, without checking whether the convexity assumptions are satisfied (and often they are not), taking the existence of solutions -to SPP- for granted. In fact, if a solution to SPP exists, our results guarantee that such a solution can be formulated recursively in term of the SPFE and, of course, this also provides a solution to PP.

[34]It can be shown that, under some restrictive conditions on utility functions and productions functions, the feasible set is convex. For example, $u(c) \equiv e^{-\gamma c}$; $c_t^g \leq \overline{c}^g$ for $\overline{c}^g$ small, $F''' - c^g \frac{\partial^2(F'/F)}{\partial^2 k} < 0$ for all $c^g \leq \overline{c}^g$, and θ_t=1 guarantees concavity of $h(\cdot,\cdot,\cdot,\cdot,s)$, a much stronger condition than the quasi-concavity assumption of **B1b**.

REFERENCES

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

Alvarez, F. and N. L. Stokey (1995): *Dynamic Programming with Homogeneous Functions.* Unpublished manuscript, University of Chicago.

Bellman, R. (1957): *Dynamic Programming.* Princeton, N.J.: Princeton University Press.

Benhabib, J., A. Rustichini and A. Velasco (1997): "Public Spending and Optimal Taxes Without Commitment," mimeo, New York University.

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

Chang, R. (1996): "Credible Monetary Policy with Long-Lived Agents: Recursive Approaches," Working Paper 96–20, Federal Reserve Bank of Atlanta.

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

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

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

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

Green, E. J. (1987): "Lending and the Smoothing of Uninsurable Income" in *Contractual Arrangements for Intertemporal Trade* ed. by E.

C. Prescott and N. Wallace, Minneapolis??: University of Minnesota Press.

HANSEN, L. P., EPPLE, D. AND ROBERDS, W. (1985): "Linear-Quadratic Duopoly Models of Resource Depletion" in *Energy, Foresight and Strategy* ed. by T. J. Sargent, Washington, DC: Resources for the Future.

KEHOE, P. AND F. PERRY (1998): "International Business Cycles with Endogenous Market Incompleteness", working paper, Federal Reserve Bank of Minneapolis.

KLETZER, K. M. AND B. D. WRIGHT (1998): "Sovereign Debt as Intertemporal Barter," mimeo, University of California, Santa Cruz.

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 Control* 2, 79–91.

LUCAS, R. E. AND N. STOKEY (1983): "Optimal Fiscal and Monetary Policy," *Journal of Monetary Economics.*

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

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

MARCET, A., T. J. SARGENT AND J. SEPPALA (1996): "Optimal Taxation without State-Contingent Debt," Working paper, Universitat Pompeu Fabra.

MARIMON, R. AND A. SCOTT (eds.) (1998): *Computational Methods for the Study of Dynamic Economies.* Oxford, UK: Oxford University Press.

Mas-Colell, A. and W. R. Zame. (1991): "Equilibrium Theory in Infinite Dimensional Spaces," in *Handbook of Mathematical Economics, Volume IV* ed. by W. Hildenbrand and H. Sonnenschein, Elsevier Science Publishing B.V.

Neveu, J. (1970): Basis Mathematiques du Calcul des Probabilities (2nd ed.) Paris: Masson (English translation of the first ed.: 1965. San Francisco, Holden Day).

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

Rojas, G. (1993): "Optimal Taxation in a Stochastic Growth Model with Public Capital," Working Paper, Universitat Pompeu Fabra.

Rustichini, A. (1996): "Dynamic Programing Solutions of Incentive Constrained Problems," Tilburg University, CentER Disc. Paper No. 9697.

Sargent, T. J. (1999): *The Conquest of American Inflation.* (forthcoming) Princeton, NJ: Princeton University Press,

Sargent, T. J. and L. Ljungqvist (1998): *Recursive Macroeconomic Theory.* (forthcoming) Stanford University.

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

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