Repeated Moral Hazard and Recursive Lagrangeans

Antonio Mele*
Universitat Pompeu Fabra, Barcelona (Spain)

November 8th, 2008
JOB MARKET PAPER

Abstract

Dynamic agency models with endogenous state variables or many agents are numerically intractable, or at least very difficult to solve. This problem prevented researchers from analyzing an important class of economic problems, e.g. extensions of the basic unemployment insurance framework, models of international risk sharing in production economies with private information, and optimal executive compensation schemes.

In this paper, I show how to solve these models, by building a general theoretical framework for the analysis of repeated moral hazard problems. The main contribution is the extension of the recursive Lagrangean techniques à la Marcet and Marimon (2008) to models with repeated moral hazard. My method can easily deal with large state spaces and many agents, allowing the analysis of currently untractable problems.

1 Introduction

In this paper, I show how to solve repeated moral hazard models with the use of recursive Lagrangean techniques. My approach allows the analysis of models with many state variables and many agents, which are instead untractable with commonly used solution strategies. Moreover, my methodology is simpler and numerically faster than the alternatives. I present the main idea in a simple model of dynamic agency, and

*Acknowledgements: I am grateful to Albert Marcet for his suggestions and long fruitful discussions on the topic. I also owe a special thank to Luigi Balletta and Sevi Rodriguez-Mora for advices at a very early stage of the work, and to Davide Debortoli and Ricardo Nunes for their generosity in discussing infinitely many numerical aspects of the paper. This paper has benefitted from comments by Klaus Adam, Sofia Bauducco, Toni Braun, Filippo Bruttì, Andrea Caggese, Francesco Caprioli, Martina Cecioni, José Dorich, Harald Fadinger, Eva Luethi, Angelo Mele, Matthias Messner, Krisztina Molnar, Juan Pablo Nicolini, Nicola Pavoni, Josep Pijoan-Mas, Michael Reiter, Pontus Rendahl, Gilles Saint-Paul, Daniel Samano, Antonella Tutino and from participants at Macro Break and Macro Discussion Group at Universitat Pompeu Fabra, SED Meeting 2008 in Cambridge (MA), Midwest Economic Theory Meeting 2008 in Urbana-Champaign, 63rd European Meeting of the Econometric Society 2008 in Milan, 14th CEF Conference 2008 in Paris. This paper was awarded the prize of the 2008 CEF Student Contest by the Society for Computational Economics. All mistakes are mine. email: antonio.mele@upf.edu
then I show examples of economic models that are either very difficult to solve or even untractable under the traditional approach, but do not pose significant difficulties with my techniques.

There has been a lot of research, in last two decades, on dynamic versions of the principal-agent model\(^1\). Typically these models do not have closed form solution, therefore it is necessary to solve them numerically. The main technical difficulty is that the optimal allocation is history-dependent: the principal must keep track of the whole history of shock realizations, use it to extract information about the agent’s unobservable behavior, and reward or punish the agent accordingly. As a consequence, it is not possible to derive a standard recursive representation of the principal’s intertemporal maximization problem. The traditional way of dealing with this complication is based on the promised utilities approach: the model can be transformed in an auxiliary problem with the same solution, in which the principal optimally chooses allocations and agent’s continuation value, taking as given the continuation value chosen in the previous period. The latter (also called promised utility) incorporates the whole history of the game, and hence continuation value becomes a new endogenous state variable to be chosen optimally. By using a standard argument, due to Abreu, Pearce and Stacchetti (1990) (APS henceforth) among others, it can be shown that the auxiliary problem has a recursive representation in a new state space that includes the continuation value and the state variables of the original problem. However, there is an additional complication: promised utilities must belong to a feasible set, which has to be characterized numerically before computation of the optimal allocation\(^2\). It is easy to characterize this set if there is just one exogenous shock, but it becomes complicated, if not computationally impossible, in models with several endogenous states. Therefore, with state-of-the-art approach, there is a huge class of models that are untractable even with numerical methods.

My paper provides a way to overcome the limits of the promised utilities approach, by extending the techniques of recursive Lagrangeans developed in Marcet and Marimon (2008) (MM henceforth) to the dynamic agency model. With respect to the traditional approach, the main gain is in terms of tractability: under MM, I do not have to characterize any feasible set, since the recursive representation of the principal-agent problem is always well-defined with no need for additional constraints. Therefore, it is possible to find a solution also in presence of several endogenous state variables and many agents. I provide an algorithm based on the recursive Lagrangean which is much faster than the usual dynamic programming techniques.

---

\(^1\)Recent contributions have focused both on the case in which agent’s consumption is observable (see for example Rogerson (1985a), Spear and Srivastava (1987), Thomas and Worrall (1990), Phelan and Townsend (1991), Fernandes and Phelan (2000)) and more recently on the case in which agents can secretly save and borrow (Werning (2001), Abraham and Pavoni (2006, 2008, forthcoming)); other works have explored what happens with the presence of more than one agent (see e.g. Zhao (2007) and Friedman (1998)), while few researchers have extended the setup to production economies with capital (Clementi, Cooley and Di Giannatale (2008a,2008b)). Among applications, a non-exhaustive list includes unemployment insurance (Hopenhayn and Nicolini (1997), Shimer and Werning (forthcoming), Werning (2002), Pavoni (2007, forthcoming)), executive compensation (Clementi et al. (2008a,2008b), Clementi et al. (2006), Atkeson and Cole (2008)), entrepreneurship (Quadrini (2004), Paulson et al. (2006)), credit markets (Lehnert et al. (1999), and many more.

\(^2\)The feasible set is the fixed point of a set-operator (see Abreu, Pierce and Stacchetti (1990) for details). The standard numerical algorithm starts with an initial set large enough, and iteratively converges to the fixed point. Sleet and Yeltekin (2003) provide an efficient way of computing the value correspondence approximation.
and does not suffer from the same dimensionality issues.

To illustrate the method, I first apply it to the simplest version of the dynamic agency model as in Spear and Srivastava (1987). I use a first-order approach: I solve the agent’s problem by taking first-order conditions with respect to effort, and using them as constraints in the principal’s maximization problem. I write down the Lagrangean, and using arguments similar to MM, I show that this saddle-point problem is recursive in an enlarged state space, which includes the stochastic output and an endogenously evolving Pareto-Negishi weight attached to agent’s utility. The latter has a natural interpretation: it summarizes the principal’s promises, according to which the agent is rewarded or punished. If a "good" realization of the output is observed, the Pareto-Negishi weight increases, therefore the principal cares more about the utility of the agent and the agent gets more consumption than in the previous period; analogously, if a "bad" outcome happens, the Pareto-Negishi weight decreases, hence the principal cares less about the utility of the agent and accordingly the agent gets less consumption than in the previous period. With the optimal choice of the Pareto-Negishi weight, the principal guarantees that the optimal allocation is incentive compatible.

Finally, I can obtain a solution from the Lagrangean first-order conditions. This methodology is much simpler to implement, and less mathematically and computationally demanding than APS techniques.

Extending the basic approach to models with several state variables is straightforward. Imagine, to fix ideas, that we want to modify the baseline dynamic principal-agent model, by introducing observable capital accumulation. Assume that output is produced through a production technology that uses capital, and it is affected by a productivity shock, the distribution of which depend on agent’s effort. It is easy to show that the Lagrangean associated with this extended model is recursive in a state space that includes the productivity shock, the Pareto-Negishi weight and the capital stock.

The numerical algorithm builds on the previous theoretical framework. The basic idea is to find approximated policy functions by solving Lagrangean first-order conditions. This algorithm is extremely fast in comparison with APS techniques. Computational speed depends in part on the fact that there is no need to characterize the feasible set for promised utilities. However, the main gain is obtained because solving a nonlinear system of equations is much faster than value function iteration.

After the detailed characterization of the optimal allocation in a model without endogenous states, I present few examples which are thought by the profession to be difficult to solve: I show how to use the recursive Lagrangean approach in a repeated moral hazard setup where the agent can accumulate assets.

---

3In order to be sure that agent’s first-order conditions are sufficient to get the optimal solution for the problem of the principal, I assume that Rogerson (1985b) conditions of monotone likelihood ratio and convex distribution function are satisfied.

4Second-order conditions can be an issue in these models. The researcher can control for this problem by starting from different initial conditions and checking if the algorithm always converges to the same solution. All examples presented in my paper are robust to this check.

5The procedure is an application of the collocation method (see Judd (1998)): first, approximate the policy functions for allocations, the agent’s continuation value and the value of the problem, over a set of grid nodes, with standard interpolation techniques (cubic splines or Chebichev polynomials); then, solve the Lagrangean first-order conditions with a nonlinear equation solver. Details are provided in the next sections.
without being monitored by the principal (as in Werning (2001) and Abraham and Pavoni (2006, forthcoming)), and in a dynamic risk-sharing problem with several agents where output depends on unobservable effort (as in Zhao (2007) and Friedman (1998)\(^6\)). In both frameworks, I obtain a recursive representation in the state space that contains the natural states and the endogenous Pareto-Negishi weight(s).

Finally, I present three examples of models that are untractable under the APS approach: a model of unemployment insurance with hidden assets and human capital, a problem of optimal executive compensation scheme, and an international risk-sharing model with moral hazard and physical capital.

The paper is organized as follows: Section 2 introduces the basic framework of repeated moral hazard and explains how to obtain the recursive Lagrangean. Section 3 shows examples of models that are tractable with APS techniques but are very difficult to solve numerically. Section 4 provides economic models that are untractable with APS techniques, and shows how to solve them with the Lagrangean approach. Section 5 explains the details of the algorithm, and provides some numerical simulation for the models described in previous sections. Section 6 concludes.

2 The basic model

In order to illustrate the Lagrangean approach, I start with a dynamic agency problem without endogenous states, where APS do not pose significant problems. In the next sections, I will extend the analysis to other setups in which the presence of endogenous state variables makes the use of APS techniques challenging for the researchers.

The economy is inhabited by a risk neutral principal and a risk averse agent. Time is discrete, and the state of the world follows an observable Markov process \( \{ s_t \}_{t=0}^{\infty} \), where \( s_t \in S \), and \( \#S = I \). The realizations of the process are public information. I will denote with subscripts the single realizations, and with superscripts the histories:

\[ s^t \equiv \{ s_0, ..., s_t \} \in S^{t+1} \]

At each period, the agent gets a state-contingent income flow \( y(s_t) \), enjoys consumption \( c_t(s^t) \), receives a transfer \( \tau_t(s^t) \) from the principal, and exerts a costly unobservable action \( a_t(s^t) \in A \subseteq \mathbb{R}^+ \), \( \{ 0 \} \in A \), and \( A \) is bounded. I will refer to \( a_t(s^t) \) as action or effort.

The costly action affects the future probability distribution of the state of the world. For simplicity, let \( \tilde{s}_i, i = 1, 2, ..., I \) be the possible realizations of \( \{ s_t \} \) and let them be ordered such that \( y(s_t = \tilde{s}_1) < y(s_t = \tilde{s}_2) < ... < y(s_t = \tilde{s}_I) \). Let \( \pi(s_{t+1} = \tilde{s}_i \mid s_t, a_t(s^t)) \) be the probability that state tomorrow is \( \tilde{s}_i \in S \) conditional on

\(^6\)Friedman (1998) uses a very similar approach to the one presented in my paper. He analyzes a dynamic risk sharing problem in which there is a finite number of agents, and each of them exerts unobservable effort (this model is briefly presented in Section 3). As in my work, he characterizes the recursivity of the optimal contract by using Pareto-Negishi weights instead of continuation values, also if he does not directly apply the Lagrangean approach. His work is focused on theoretical results, though, and it does not provide any numerical example. Finally, he does not exploit the homogeneity properties of the value function to reduce the dimensionality of the state space, as I do here.
past state and effort exerted by the agent at the beginning of the period\(^7\), with \(\pi (s_0 = \hat{s}_t) = 1\). I assume \(\pi (\cdot)\) is twice continuously differentiable in \(a_t (s^t)\), and has full support: \(\pi (s_{t+1} = \hat{s}_t | s_t, a) > 0 \; \forall i, \forall a, \forall s_t\). Let \(\Pi \left( s^{t+1} | s_0, a^t (s^t) \right) = \prod_{j=0}^{t} \pi \left( s_{j+1} | s_j, a_j (s^t) \right)\) be the probability of history \(s^{t+1}\) induced by the history of unobserved actions \(a^t (s^t) \equiv (a_0 (s^0), a_1 (s^1), \ldots, a_t (s^t))\).

The instantaneous utility of the agents is

\[
u \left( c_t \left( s^t \right) \right) - v \left( a_t \left( s^t \right) \right)
\]

with \(u (\cdot)\) strictly increasing, strictly concave and satisfying Inada conditions, while \(v (\cdot)\) is strictly increasing and strictly convex; both are twice continuously differentiable. Agents do not accumulate assets autonomously: the only source of insurance is the principal. Then, the budget constraint of an agent will be simply:

\[
c_t \left( s^t \right) = y (s_t) + \tau_t \left( s^t \right) \quad \forall s^t, t \geq 0
\]

Both principal and agent are fully committed once they sign the contract at time zero.

A contract (or allocation) in this framework is a plan \((a^\infty, c^\infty, \tau^\infty) \equiv \left\{ a_t (s^t), c_t (s^t), \tau_t (s^t) \right\} \quad \forall s^t \in S^{t+1} \}_{t=0}^{\infty}\) that belongs to the following set:

\[
\Gamma^{MH} \equiv \left\{ (a^\infty, c^\infty, \tau^\infty) : a_t (s^t) \in A, \quad c_t (s^t) \geq 0, \quad \tau_t (s^t) = c_t (s^t) - y (s_t) \quad \forall s^t \in S^{t+1}, t \geq 0 \right\}
\]

Assume, for simplicity, that the discount factor of the agent and the principal is the same. The principal evaluates allocations according to the following

\[
P (s_0; a^\infty, c^\infty, \tau^\infty) = -\sum_{t=0}^{\infty} \sum_{s^t} \beta^t \tau_t (s^t) \Pi \left( s^t | s_0, a^{t-1} (s^{t-1}) \right) \quad (1)
\]

\[
= \sum_{t=0}^{\infty} \sum_{s^t} \beta^t \left[ y (s_t) - c_t (s^t) \right] \Pi \left( s^t | s_0, a^{t-1} (s^{t-1}) \right)
\]

therefore efficient contracts can be characterized by maximizing (1), subject to incentive compatibility and to the requirement of providing at least a minimum level of ex-ante utility \(V^{out}\) to the agent:

\[
W (s_0) = \max_{\left\{ a_t (s^t), c_t (s^t) \right\}_{t=0}^{\infty} \in \Gamma^{MH}} \sum_{t=0}^{\infty} \sum_{s^t} \beta^t \left[ y (s_t) - c_t (s^t) \right] \Pi \left( s^t | s_0, a^{t-1} (s^{t-1}) \right) \quad (2)
\]

\[
s.t. \quad a^\infty \in \arg \max_{s_0 \in (a_t (s^t))_{t=0}^{\infty}} \sum_{t=0}^{\infty} \sum_{s^t} \beta^t \left[ u \left( c_t (s^t) \right) - v \left( a_t (s^t) \right) \right] \Pi \left( s^t | s_0, a^{t-1} (s^{t-1}) \right) \geq V^{out} \quad (3)
\]

I will call this the original problem. Notice that (2) is a very complicated object. In this work, I use the first order conditions of the agent’s problem as a substitute for the constraint (2). In order to guarantee

---

\(^7\) Notice that I allow for persistence; in the numerical examples, I focus on i.i.d. shocks, but it should be clear that persistence does not create particular problems neither theoretically nor numerically.
that this substitution leads to the actual solution of the original problem, I assume that Rogerson (1985b) conditions of monotone likelihood ratio and convexity of the distribution are satisfied:

**Condition 1 (Monotone Likelihood-Ratio Condition (MLRC))** \( \hat{a} \leq \tilde{a} \implies \frac{\pi(s_{t+1} = s_i | s_t, \hat{a})}{\pi(s_{t+1} = s_i | s_t, \tilde{a})} \) is nonincreasing in \( i \).

The above property can be restated in a simpler way: if \( \pi(\cdot) \) is differentiable, then MLRC is equivalent to \( \frac{\pi_{a}(s_{t+1} = s_i | s_t, a)}{\pi_{a}(s_{t+1} = s_i | s_t, \hat{a})} \) being nondecreasing in \( i \) for any \( a \), where \( \pi_{a}(s_{t+1} = s_i | s_t, a) \) is the derivative of \( \pi(\cdot) \) with respect to \( a \). An important consequence of the MLRC is the following: let \( F(\cdot) \) be the cumulative distribution function of \( \pi(\cdot) \); then MLRC implies that the density function \( F'(s_{t+1} = s_i | s_t, a) \) is nonpositive for any \( i \) and every \( a \). Therefore, more effort implies a first order stochastic dominance shift of the distribution (see Rogerson (1985b)).

**Condition 2 (Convexity of the Distribution Function Condition (CDFC))** \( F''(s_{t+1} = s_i | s_t, a) \) is nonnegative for any \( i \) and every \( a \).

This condition implies that the cumulative distribution function is convex.

In more intuitive terms, MLRC asks for the state of nature to be "sufficiently informative" about the unobservable effort, while CDFC says that this informativeness has "decreasing returns to scale".

I now define the problem of the agent and I derive his first order conditions with respect to effort. The problem of the agent, given the principal’s strategy profile \( \tau^\infty \equiv \{\tau_t(s^t)\}_{t=0}^\infty \), is:

\[
V(s_0; \tau^\infty) = \max_{\{c_t(s^t), a_t(s^t)\}_{t=0}^\infty} \left\{ \sum_{t=0}^{\infty} \sum_{s^t} \beta^t \left[ u(c_t(s^t)) - \Phi \Pi(s^t | s_0, a^{t-1}(s^{t-1})) \right] \right\}
\]

The first order condition for effort is:

\[
u'(a_t(s^t)) = \sum_{j=1}^{\infty} \beta^j \sum_{s^{t+j} | s^t} \pi_{a_t}(s_{t+1} = s_i | s_t, a_t(s^t)) \times [u(c_{t+j}(s^{t+j})) - \Phi \Pi(s^{t+j} | s^{t+1}, a^{t+j}(s^{t+j}))]
\]

Intuitively, the marginal cost of effort today (LHS) has to be equal to future expected benefits (RHS) in terms of expected future utility. The use of (4) is key to my approach, since it allows me to write the Lagrangean of the principal’s problem. In the following, for simplicity I will refer to (4) as the *incentive-compatibility constraint* (ICC).
We can write the Pareto problem of the principal as:

\[
W(s_0) = \max_{\{a_t(s^t), c_t(s^t)\}_{t=0}^{\infty} \in G_{MR}} \sum_{t=0}^{\infty} \sum_{s^t} \beta^t \left[ y(s_t) - c_t(s^t) \right] \Pi(s^t | s_0, a^{t-1}(s^{t-1})) \\
\text{s.t. } \left. u'(a_t(s^t)) = \sum_{j=1}^{\infty} \beta^j \sum_{s^{j+t+j}} \frac{\pi_a(s_{t+1} | s_t, a_t(s^t))}{\pi(s_{t+1} | s_t, a_t(s^t))} \times \right. \\
\times [u(c_{t+j}(s^{j+t})) - u(a_{t+j}(s^{j+t}))] \Pi(s^{j+t} | s^t, a^{t+j-1}(s^{t+j-1} | s^t)) \\
\forall s^t, t \geq 0
\]

\[
\sum_{t=0}^{\infty} \sum_{s^t} \beta^t [u(c_t(s^t)) - v(a_t(s^t))] \Pi(s^t | s_0, a^{t-1}(s^{t-1})) \geq V^{out}
\]

2.1 The Lagrangean approach

In order to write the Lagrangean of the Pareto problem, notice that (3) must be binding in the optimum: otherwise, the principal can increase her expected discounted utility by asking the agent to increase effort in period 0 by \(\delta > 0\), provided that \(\delta\) is small enough. Therefore (3) will be associated with a strictly positive Lagrange multiplier (say, \(\gamma\)), which will be a function of \(V^{out}\): for every \(V^{out}\), there will be a \(\gamma\) associated with (3). This Lagrange multiplier can be seen as a Pareto-Negishi weight on the agent’s utility. Since each \(\gamma\) implies a unique \(V^{out}\), I can fully characterize the Pareto frontier of this economy by solving the problem for different values of \(\gamma\) between zero and infinity. Hence, in the following, I am going to consider \(\gamma\) as a parameter, that represents the constraint (3). Moreover, notice that by fixing \(\gamma\), \(V^{out}\) will appear in the Lagrangean only in the constant term \(\gamma V^{out}\), thus it will be irrelevant for the optimal allocation. Given these considerations, Problem (PP1) can be seen as the constrained maximization of a social welfare function, where the Pareto weight for the principal and the agent are, respectively, 1 and \(\gamma\):

\[
W^{SWF}(s_0) = \max_{\{a_t(s^t), c_t(s^t)\}_{t=0}^{\infty} \in G_{MR}} \sum_{t=0}^{\infty} \sum_{s^t} \beta^t \left[ y(s_t) - c_t(s^t) \right] \Pi(s^t | s_0, a^{t-1}(s^{t-1})) + \\
+ \gamma \sum_{t=0}^{\infty} \sum_{s^t} \beta^t [u(c_t(s^t)) - v(a_t(s^t))] \Pi(s^t | s_0, a^{t-1}(s^{t-1}))
\]

s.t. \(u'(a_t(s^t)) = \sum_{j=1}^{\infty} \beta^j \sum_{s^{j+t+j}} \frac{\pi_a(s_{t+1} | s_t, a_t(s^t))}{\pi(s_{t+1} | s_t, a_t(s^t))} \times \)
\times [u(c_{t+j}(s^{j+t})) - u(a_{t+j}(s^{j+t}))] \Pi(s^{j+t} | s^t, a^{t+j-1}(s^{t+j-1} | s^t))

Let \(\beta^t \pi_t(s^t) \Pi(s^t | s_0, a^{t-1}(s^{t-1}))\) be the Lagrange multiplier associated to each ICC. I can therefore write the Lagrangean as:

\[
L(s_0, \gamma, c^\infty, a^\infty, \lambda^\infty) = \sum_{t=0}^{\infty} \sum_{s^t} \beta^t \left[ y(s_t) - c_t(s^t) + \gamma [u(c_t(s^t)) - v(a_t(s^t))] \right] \Pi(s^t | s_0, a^{t-1}(s^{t-1})) + \\
- \sum_{t=0}^{\infty} \sum_{s^t} \beta^t \pi_t(s^t) \left\{ u'(a_t(s^t)) - \sum_{j=1}^{\infty} \beta^j \frac{\pi_a(s_{t+1} | s_t, a_t(s^t))}{\pi(s_{t+1} | s_t, a_t(s^t))} \times \right. \\
\times [u(c_{t+j}(s^{j+t})) - u(a_{t+j}(s^{j+t}))] \Pi(s^{j+t} | s^t, a^{t+j-1}(s^{t+j-1} | s^t)) \right\} \Pi(s^t | s_0, a^{t-1}(s^{t-1}))
\]

7
The Lagrangean can be manipulated with simple algebra to get the following expression:

\[ L(s_0, \gamma, c^\infty, a^\infty, \lambda^\infty) = \sum_{t=0}^{\infty} \sum_s \beta^t \left\{ y(s_t) - c_t(s^t) + \phi_t(s^t) \left[ u(c_t(s^t)) - v(a_t(s^t)) \right] \right\} + \\
- \lambda_t(s^t) v'(a_t(s^t)) \} \Pi(s^t | s_0, a^{t-1}(s^{t-1})) \]

where

\[ \phi_t(s^{t-1}, s_t) = \gamma + \sum_{i=0}^{t-1} \lambda_i(s^i) \frac{\pi_a(s_{i+1} | a_i(s^i))}{\pi(s_{i+1} | s_i, a_i(s^i))} \]

The intuition is simple. For any \( s^t \), the expression \( \lambda_t(s^t) \frac{\pi_a(s_{t+1} | s_t, a_t(s^t))}{\pi(s_{t+1} | s_t, a_t(s^t))} \) is a planner’s promise about how much she will increase the weight of agent’s welfare in the future, depending on which realization of state \( s_{t+1} \) is observed. By keeping track of all \( \lambda \)'s and \( \frac{\pi_a}{\pi} \)'s realized in the past, \( \phi_t(s^t) \) summarizes all the promises made by the planner in previous periods. In this framework, there is straightforward interpretation of \( \phi_t(s^t) \): it is the Pareto-Negishi weight of the agent’s lifetime utility, that evolves endogenously in order to track agent’s effort with the following recursive law of motion:

\[ \phi_{t+1}(s^t, s) = \phi_t(s^t) + \lambda_t(s^t) \frac{\pi_a(s_{t+1} \in s | s_t, a_t(s^t))}{\pi(s_{t+1} \in s | s_t, a_t(s^t))} \quad \forall s \in S \]

(5)

\[ \phi_0(s^0) = \gamma \]

To better understand the role of \( \phi_t(s^t) \), let us assume there are only two possible realizations of the state of nature: \( s_t \in \{s_L, s_H\} \). At time 0, the weight is equal to \( \gamma \). In period 1, given our assumption on the likelihood ratio, the Pareto-Negishi weight is higher than \( \gamma \) if the principal observes \( s_H \), while it is lower than \( \gamma \) if she observes \( s_L \) (a formal proof of this fact is obtained in Lemma 1). Therefore the agent is rewarded by a higher weight in the social welfare function of the principal (i.e., the principal cares more about him) if a good state of nature is observed, while it is punished by a lower weight (i.e., the principal cares less about him) if a bad state of nature happens.

### 2.2 Recursive formulation

By the duality theory (see for example Luenberger (1969)), we know that a solution of the original problem corresponds to a saddle point of the Lagrangean, i.e. the contract

\[ (c^{\infty*}, a^{\infty*}, x^{\infty*}) = \left\{ c^*_t(s^t), a^*_t(s^t), y(s_t) - c^*_t(s^t) \right\} \quad \forall s^t \in S^{t+1}_{t=0} \]

is a solution for the original problem if there exist a sequence \( \{\lambda^*_t(s^t) \} \quad \forall s^t \in S^{t+1}_{t=0} \) of Lagrange multipliers such that \((c^{\infty*}, a^{\infty*}, \lambda^{\infty*}) = (c^*_t(s^t), a^*_t(s^t), \lambda^*_t(s^t)) \quad \forall s^t \in S^{t+1}_{t=0} \) satisfy:

\[ L(s_0, \gamma, c^{\infty*}, a^{\infty*}, \lambda^{\infty*}) \leq L(s_0, \gamma, c^{\infty*}, a^{\infty*}, \lambda^{\infty*}) \leq L(s_0, \gamma, c^{\infty*}, a^{\infty*}, \lambda^{\infty*}) \]

It is possible to recursively characterize the solutions of the Lagrangean. In particular, it is possible to show that value and policy functions depend on the state of the world \( s_t \) and the Pareto-Negishi weight \( \phi_t(s^t) \).
The reader not interested in the details can skip this section and jump directly to the characterization of the optimal allocation.

I follow the strategy of MM by showing that a generalized version of (PP 1) is recursive in an enlarged state space. Let me define the following generalized version of (PP 1):

\[
W_0^{SWF}(s_0) = \max_{(a_t(s_t), c_t(s_t)\mid t=0} \sum_{t=0}^{\infty} \sum_{s_t} \beta^t \left[ y(s_t) - c_t(s_t) \right] \Pi(s^t \mid s_0, a^{t-1}(s^{t-1})) + \\
\gamma \sum_{t=0}^{\infty} \sum_{s_t} \beta^t \left( u(c_t(s_t)) - v(a_t(s_t)) \right) \Pi(s^t \mid s_0, a^{t-1}(s^{t-1}))
\]

s.t. \( v'(a_t(s^t)) = \sum_{j=1}^{\infty} \beta^j \sum_{s^{t+j}} \pi_a(s_{t+1} \mid s_t, a_t(s_t)) \times \\
\left[ u(c_{t+j}(s^{t+j})) - v(a_{t+j}(s^{t+j})) \right] \Pi(s^{t+j} \mid s^t, a^{t+j-1}(s^{t+j-1} \mid s^t)) \forall s^t, t \geq 0
\]

Notice that if \( \beta^0 = 1 \), then we are back to (PP 1). We can write down the Lagrangean of this problem by assigning a Lagrange multiplier \( \beta^t \lambda_t(s^t) \) to each ICC constraint:

\[
L_\theta(s_0, \gamma, c^\infty, a^\infty, \lambda^\infty) = \sum_{t=0}^{\infty} \sum_{s^t} \beta^t \left[ \varphi^0 [y(s_t) - c_t(s_t)] + \gamma [u(c_t(s_t)) - v(a_t(s_t))] \right] \Pi(s^t \mid s_0, a^{t-1}(s^{t-1})) + \\
- \sum_{t=0}^{\infty} \sum_{s^t} \beta^t \lambda_t(s^t) \left[ v'(a_t(s^t)) - \sum_{j=1}^{\infty} \beta^j \sum_{s^{t+j}} \pi_a(s_{t+1} \mid s_t, a_t(s_t)) \times \\
\left[ u(c_{t+j}(s^{t+j})) - v(a_{t+j}(s^{t+j})) \right] \Pi(s^{t+j} \mid s^t, a^{t+j-1}(s^{t+j-1} \mid s^t)) \right] \Pi(s^t \mid s_0, a^{t-1}(s^{t-1}))
\]

Notice that \( r(a, c, s) \equiv y(s) - c \) is uniformly bounded by natural debt limits, so there exists a lower bound \( \kappa \) such that \( r(a, c, s) \geq \kappa \). We can therefore define \( \kappa < \frac{\phi}{1-\beta} \). Define \( \varphi(\phi, \lambda, s^t) \equiv \phi + \lambda \pi_a(s^t \mid s_0, a) \), \( h_0^P(a, c, s) \equiv r(a, c, s), h_0^{ICC}(a, c, s) \equiv u(c) - v(a), h_1^{ICC}(a, c, s) \equiv -v'(a), \theta \equiv \left[ \begin{array}{c} \phi^0 \phi \\ \varphi \end{array} \right] \in R^2, \chi \equiv \left[ \begin{array}{c} \lambda^0 \lambda \end{array} \right] \) and

\[
h(a, c, \theta, \chi, s) \equiv \theta h_0(a, c, s) + \chi h_1(a, c, s) = \left[ \begin{array}{c} \phi^0 \phi \\ \varphi \end{array} \right] \left[ \begin{array}{c} h_0^P(a, c, s) \\ h_0^{ICC}(a, c, s) \end{array} \right] + \left[ \begin{array}{c} \lambda^0 \lambda \end{array} \right] \left[ \begin{array}{c} h_1^P(a, c, s) \\ h_1^{ICC}(a, c, s) \end{array} \right]
\]

which is homogenous of degree 1 in \((\theta, \chi)\). The Lagrangean can be written as:

\[
L_\theta(s_0, \gamma, c^\infty, a^\infty, \lambda^\infty) = \sum_{t=0}^{\infty} \sum_{s^t} \beta^t h(a_t(s^t), c_t(s^t), \theta_t(s^t), \chi_t(s^t), s_t) \Pi(s^t \mid s_0, a^{t-1}(s^{t-1}))
\]

where

\[
\theta_{t+1}(s^t, \tilde{s}) = \varphi(\theta_t(s^t), \chi_t(s^t), \tilde{s}) \forall \tilde{s} \in S
\]

\[
\theta_0(s^0) = \left[ \begin{array}{c} \phi^0 \phi \\ \chi \end{array} \right]
\]
Notice that the constraint defined by \( h_P^1 (a, c, s) \) is never binding by definition, therefore \( \lambda_0^0 (s^t) = 0 \) and \( \phi_0^0 (s^t) = \phi \) \( \forall s^t, t \geq 0 \), which implies that the only relevant state variable is \( \phi_t (s^t) \). We can associate a saddle point functional equation to this Lagrangean

\[
J (s, \theta) = \min_{\chi > 0} \max_{a, c} \left\{ h (a, c, \theta, s) + \beta \sum_{s'} \pi (s' \mid s, a) J (s', \theta' (s')) \right\} \tag{SPFE}
\]

subject to \( \theta' (s^t) = \theta + \chi \pi_a (s' \mid s, a) \pi (s' \mid s, a) \) \( \forall s'^t \).

In order to show that there is a unique value function \( J (s, \theta) \) that solves (6), it is sufficient to prove that the operator on the right hand side of the functional equation is a contraction\(^8\).

There are two technical differences with the original framework in MM. First, the endogenous evolution of the Pareto-Negishi weight is a deviation from MM, since in their paper the law of motion of the costate variable \( \theta_t (s^t) \) only depends on \( \chi_t (s^t) \), while here also depends on \( a_t (s^t) \). Second, the probability distribution of the future states is endogenous and depends on the optimal effort \( a_t (s^t) \). I show in Proposition 1 that the argument in MM works also here with some minor modifications.

**Proposition 1** Fix an arbitrary constant \( K > 0 \) and let \( K_\theta = \max \{ K, K \| \theta \| \} \). The operator

\[
(T_K f) (s, \theta) \equiv \min_{\chi > 0 \| \chi \| \leq K_\theta} \max_{a, c} \left\{ h (a, c, \theta, s) + \beta \sum_{s'} \pi (s' \mid s, a) f (s', \theta' (s')) \right\}
\]

subject to \( \theta' (s') = \theta + \chi \pi_a (s' \mid s, a) \pi (s' \mid s, a) \) \( \forall s' \)

is a contraction.

**Proof.** Appendix A. ■

Proposition 1 shows that the saddle point problem is recursive in the state space \( (s, \theta) \in S \times \mathbb{R}^2 \). All the other theorems in MM apply directly to my framework: the result of Proposition 1 is valid for any \( K > 0 \), and since, whenever the Lagrangean has a solution the Lagrange multipliers are bounded, then a recursive solution of the Problem (SPFE) is a solution of the Lagrangean, and more importantly it is a solution of the original problem. As a consequence, we can restrict the search of optimal contracts to the set of policy functions that are Markovian in the space \( (s, \theta) \in S \times \mathbb{R}^2 \). But remember that the first element of \( \theta \) is constant for any \( t \) and the only actual endogenous state is \( \phi_t (s^t) \); therefore, from this point of view, finding the optimal contract has the same numerical complexity as finding the optimal allocations in a standard stochastic neoclassical growth model.

Notice that, since in the Lagrangean formulation we eliminated the constant \( \gamma V^{out} \), the value of the original problem is:

\[
W (s_0) = W^{SWF} (s_0) - \gamma V^{out} = J (s_0, [1 \quad \gamma ] ) - \gamma V^{out}
\]

\(^8\)Messner and Pavoni (2004) show that, also if the value of the problem 6 is unique, the policy function associated with it can be suboptimal or even unfeasible. To avoid these issues, though, it is sufficient to impose that the policy function satisfies all the constraints of the original problem. Since I solve for the Lagrangean first-order conditions, I always impose all the constraints.
where \(V^\text{out} = V(s_0; \tau^{\infty})\) is the agent’s lifetime utility implied by the optimal contract. Another important consequence of Proposition 1 is that the value function \(J(s, \theta)\) is homogeneous of degree 1 (and consequently, policy functions for allocations are homogeneous of degree zero)\(^9\): this fact will be important in the last example in Section 3.2.

### 2.3 Characterization of the optimal contract

In this section I show few properties of the optimal contract. Those properties are the analogous, in the Lagrangean approach, of well known results in the literature. Let us go back to the problem with \(\phi^0 = 1\).

We can take the first order conditions of the Lagrangean:

\[
c_t(s^t) : 0 = -1 + \phi_t(s^t) u_c(c_t(s^t))
\]

\[
a_t(s^t) : 0 = -\lambda_t(s^t) v''(a_t(s^t)) - \phi_t(s^t) v'(a_t(s^t)) +
\]

\[
+ \sum_{j=1}^{\infty} \beta^j \sum_{s^{t+j} \mid s^t} \frac{\pi_a(s_{t+1} \mid s_t, a_t(s^t))}{\pi(s_{t+1} \mid s_t, a_t(s^t))} \left\{ y(s_t) - c_t(s^t) - \lambda_{t+j}(s_{t+j}) v'(a_{t+j}(s^{t+j})) -\right.
\]

\[
\left. + \phi_{t+j}(s^{t+j}) [u(c_{t+j}(s^{t+j})) - v(a_{t+j}(s^{t+j}))] \right\} \Pi(s^{t+j} \mid s^t, a^{t+j-1}(s^{t+j-1})) +
\]

\[
+ \beta \lambda_t(s^t) \sum_{s^{t+j} \mid s^t} \frac{\partial}{\partial a} \left( \pi_a(s_{t+1} \mid s_t, a_t(s^t)) \right) [u(c_{t+1}(s^{t+1})) - v(a_{t+1}(s^{t+1}))] \pi(s_{t+1} \mid s_t, a_t(s^t))
\]

and

\[
\lambda_t(s^t) : 0 = -v'(a_t(s^t)) + \sum_{j=1}^{\infty} \sum_{s^{t+j} \mid s^t} \frac{\pi_a(s_{t+1} \mid s_t, a_t(s^t))}{\pi(s_{t+1} \mid s_t, a_t(s^t))} \times
\]

\[
\times \left[ \beta^j [u(c_{t+j}(s^{t+j})) - v(a_{t+j}(s^{t+j}))] \Pi(s^{t+j} \mid s^t, a^{t+j-1}(s^{t+j-1} \mid s^t)) \right]
\]

Lemma 1 makes clear how \(\phi_t(s^t)\) incorporates the promises of the principal. From (6) we can see that \(c_{t+1}(s^{t+1}) = u_c^{-1} \left( \frac{1}{\phi_{t+1}(s^{t+1})} \right)\), then \(c_{t+1}(s^{t+1})\) is increasing in \(\phi_{t+1}(s^{t+1})\). Lemma 1 says that, tomorrow, the principal will reward a high income realization with higher consumption than today, and a low income realization with lower consumption than today\(^10\).

**Lemma 1** In the optimal contract, \(\phi_{t+1}(s^t, s_1^t) < \phi_t(s^t) < \phi_{t+1}(s^t, \hat{s}_1^t)\) for any \(t\).

**Proof.** Appendix A. \(\blacksquare\)

The following Proposition characterizes the long run properties of the Pareto Negishi weight.

**Proposition 2** \(\phi_t(s^t)\) is a martingale that converges to zero almost surely.

**Proof.** Appendix A. \(\blacksquare\)

---

\(^9\)This is made clear in the proof.

\(^10\)Thomas and Worrall (1990) show the analogous property with APS techniques.
Proposition 2 is the well known result that $\frac{1}{u(c_{t+1}(s^{t+1}))}$ evolves as a martingale (see Rogerson (1985a)). The a.s.-convergence to zero is the so called \textit{immiseration property} that implies zero consumption almost surely as $t \to \infty$, which is a standard result in models with asymmetric information (see Thomas and Worrall (1990), for example). In this framework, the immiseration property has an intuitive interpretation: in order to keep strong incentives for the agent, the planner must ensure that the Pareto-Negishi weight goes to zero almost surely as $t \to \infty$ for any possible sequence of realizations of the income shock.

The result in Proposition 2 is obtained by using the law of motion of $\phi_t(s^t)$ and (6), which yields

$$E_t^x \left[ \frac{1}{u_c(c_{t+1}(s^{t+1}))} \right] = \frac{1}{u_c(c_t(s^t))}$$

We can use Jensen’s inequality and the strict concavity of $u(\cdot)$ to get that $E_t^x \left[ u_c(c_{t+1}(s^{t+1})) \right] > u_c(c_t(s^t))$: the profile of expected consumption is decreasing across time.

3 Examples of tractable models that are difficult to solve with APS

In this section, I provide two examples of models that are tractable with promised utilities approach, but are computationally intense. The first example is a model of repeated moral hazard with hidden assets, as in Abraham and Pavoni (2006, 2008, forthcoming) and Werning (2001, 2002), where the agent is allowed to save at a constant gross rate of return, but bond holdings are private information. The second example is a risk sharing problem where each individual must exert unobservable effort that affect the aggregate output, as in Zhao (2007) and Friedman (1998).

3.1 Repeated moral hazard with hidden assets

Werning (2001, 2002) and Abraham and Pavoni (2006, 2008) (AP from here on) analyze a model with hidden effort and hidden assets: the agent can borrow or lend without being monitored by the principal. This problem generates a continuum of incentive constraints (for each possible income realization, there is a continuum of possible asset positions). Hence the feasible set of continuation values has infinite dimension and APS techniques cannot be used. In order to overcome this complication, they characterize the optimal contract by defining an auxiliary problem, where agent’s first-order conditions over effort and bonds are used as constraints for the principal’s problem. They show that the solution of their auxiliary problem is characterized by three state variables (income, promised utility and consumption marginal utility), and can be solved recursively by value function iteration. Abraham and Pavoni (2006, forthcoming) also provide a numerical procedure to verify if the first-order approach delivers the true incentive compatible allocation. Even if their work is big step ahead in the analysis of this class of models, the use of APS arguments makes
their numerical algorithm too slow for calibration purposes and not easily adaptable to more complicated extensions. In this section, I show how the Lagrangean approach can easily deal with this framework.

Let \( \{ b_t (s^t) \}_{t=-1} \) be a sequence of one-period bond that the agent pays 1 today, getting \( R \) tomorrow. Assume that the principal cannot monitor the bond market, so that the asset accumulation is unobservable to her. Then agent’s budget constraint becomes:

\[
c_t (s^t) + b_t (s^t) = y (s_t) + \tau_t (s^t) + R b_{t-1} (s^{t-1})
\]

while the instantaneous utility function for the agent is the same as in Section 2. We have to solve now the following agent’s problem:

\[
\tilde{V} (s_0, b_{-1} ; \tau^\infty) = \max_{\{ c_t(s_t), b_t(s_t), a_t(s_t) \}_{t=\infty} \in \Gamma_{HA}} \left\{ \sum_{t=0}^{\infty} \sum_{s^t} \beta^t \left[ u (c_t (s^t)) - v (a_t (s^t)) \right] \Pi (s^t \mid s_0, a^{t-1} (s^{t-1})) \right\}
\]

where

\[
\Gamma_{HA} \equiv \{(a^\infty, c^\infty, b^\infty, \tau^\infty) : a_t (s^t) \in A, \ c_t (s^t) \geq 0, \ c_t (s^t) + b_t (s^t) = y (s_t) + \tau_t (s^t) + R b_{t-1} (s^{t-1}) \ \forall s^t \in S^{t+1}, t \geq 0 \}
\]

Accordingly, agent’s first order conditions with respect to the unobservable variables (i.e., effort and bond holdings) are (4) and the following Euler equation:

\[
u' (c_t (s^t)) = \beta R \sum_{s_{t+1}} u' (c_{t+1} (s^t, s_{t+1})) \pi (s_{t+1} \mid s_t, a_t (s^t)) \tag{8}\]

Assume \( \beta R = 1 \) to simplify algebra. The presence of hidden assets requires (8) to be included in the set of constraints for the principal’s problem.

Let \( \beta^t \eta_t (s^t) \) be the Lagrange multiplier for (8), and \( \beta^t \lambda_t (s^t) \) the Lagrange multiplier for ICC. The Lagrangean can be manipulated to get:

\[
L (s_0, \gamma, c^\infty, a^\infty, \lambda^\infty, \eta^\infty) = \sum_{t=0}^{\infty} \sum_{s^t} \beta^t \left\{ y (s_t) - c_t (s^t) + \phi_t (s^t) \left[ u (c_t (s^t)) - v (a_t (s^t)) \right] \right\} + \left[ \phi_t (s^t) \nu' (a_t (s^t)) + [\eta_t (s^t) - \beta^{-1} \zeta_t (s^t)] \nu c_t (s^t) \right] \Pi (s^t \mid s_0, a^{t-1} (s^{t-1})) \tag{9}\]

where

\[
\phi_{t+1} (s^t, \tilde{s}) = \phi_t (s^t) + \lambda_t (s^t) \frac{\pi_a (s_{t+1} = \tilde{s} \mid s_t, a_t (s^t))}{\pi (s_{t+1} = \tilde{s} \mid s_t, a_t (s^t))} \ \forall \tilde{s} \in S \text{ and } \phi_0 (s^0) = \gamma
\]

\[
\zeta_{t+1} (s^t, \tilde{s}) = \eta_t (s^t) \ \forall \tilde{s} \in S \text{ and } \zeta_0 (s^0) = 0
\]

This problem is characterized by two costate variables: the Pareto weight \( \phi_t (s^t) \) and the new costate \( \zeta_t (s^t) \), which keeps track of the Euler equation. Using the same arguments of Proposition 1, it is possible to show that the problem is recursive in the state space that includes \( (s, \phi, \zeta) \) as states variables (see Appendix B for details).
3.2 Multiple agents

The Lagrangean method can make a big difference also in models with many agents. In this section, I present a model of dynamic risk-sharing with a finite number of agents, where each of them exerts unobservable effort. Zhao (2007) solves numerically the same model with APS techniques. The setup presented here is inspired, with minor differences, by Zhao (2007).

There are \( N \) agents indexed by \( i = 1, \ldots, N \). Each agent is subject to an observable Markov state process \( \{s_{it}\}_{t=0}^{\infty} \), where \( s_{it} \in S_i \). where \( s_{i0} \) is known, and the process is common knowledge. The process is independent across agents. Let \( S \equiv \bigotimes_{i=1}^{N} S_i \) and \( s_t \equiv \{s_1, \ldots, s_N\} \in S \) be the state of nature in the economy, let \( s_t \equiv \{s_0, \ldots, s_t\} \) be the history of these realizations. In the following, let \( x_t(s_t) \equiv (x_{1t}(s_t), \ldots, x_{Nt}(s_t)) \) for any generic variable \( x \).

Each agent exerts a costly action \( a_{it}(s_t) \), which is unobservable to other players. This action affects next period distribution of states of nature: let \( \pi(s_{i,t+1} | s_{it}, a_{it}(s_t)) \) be the probability that state is \( s_{i,t+1} \) conditional on past state and effort exerted by the agent in period \( t \). Therefore, since the processes are independent across agents, we can define \( \Pi(s_{t+1} | s_0, a_t(s_t)) = \prod_{i=1}^{N} \prod_{j=0}^{t} \pi(s_{ij}, a_{ij}(s_j)) \) to be the cumulated probability of an history \( s_{t+1} \) given the whole history of unobserved actions \( a_t(s_t) \equiv (a_0(s_0), a_1(s_1), \ldots, a_t(s_t)) \). I assume \( \pi(s_{i,t+1} | s_{it}, a_{it}(s_t)) \) is differentiable in \( a_{it}(s_t) \) as many time as necessary, and I denote its derivative with respect to \( a_{it}(s_t) \) as \( \pi_{a_{it}}(s_{i,t+1} | s_{it}, a_{it}(s_t)) \).

The utility of each agent is

\[
u(c_{it}(s_t)) - \nu(a_{it}(s_t))\]

and the resource constraint of the economy is:

\[
\sum_{i=1}^{N} c_{it}(s_t) \leq \sum_{i=1}^{N} y_{it}(s_{it})
\]

(10)

where \( y_{it}(s_{it}) \) is the stochastic endowment of each agent.

A feasible contract is a sequence \( (a^\infty, c^\infty) \equiv \{c_t(s_t), a_t(s_t)\}_{t=0}^{\infty} \) such that (10) is satisfied. Therefore

\[
\Gamma^{MA} \equiv \{(a^\infty, c^\infty) : a_t(s_t) \in A, \quad c_t(s_t) \geq 0, \quad \sum_{i=1}^{N} c_{it}(s_t) \leq \sum_{i=1}^{N} y_{it}(s_{it}) \quad \forall s_t \in S_{t+1}, t \geq 0\}
\]

Let \( \omega \equiv \{\omega_i\}_{i=1}^{N} \) be a vector of weights, and assume MLRC and CDFC are satisfied in this economy. Since first-order condition with respect to effort for each agent is the same as in Section 2, the constrained efficient
allocation is the solution of the following maximization problem:

\[
P(s_0) = \max_{\{c_t(s^t), a_t(s^t)\}_{t=0}^\infty} \left\{ \sum_{i=1}^N \sum_{t=0}^\infty \beta^t \sum_{s^t} \omega_i \left[ u(c_{it}(s^t)) - v(a_{it}(s^t)) \right] \Pi(s^t | s_0, a^{t-1}(s^{t-1})) \right\}
\]

s.t. \( u'(a_{it}(s^t)) = \sum_{j=1}^\infty \sum_{s^{t+j} \in \mathcal{S}^t} \beta^j \pi_{ai}(s_{it}, a_{it}(s^t)) \left[ u(c_{it+j}(s^{t+j})) - v(a_{it+j}(s^{t+j})) \right] \Pi(s^{t+j} | s^t, a^{t+j-1}(s^{t+j-1})) \quad \forall i = 1, ..., N \)

\[
\sum_{i=1}^N c_{it}(s^t) = \sum_{i=1}^N y_i \pi_i(s_{it})
\]

Let \( \beta^t \lambda_{it}(s^t) \) be the Lagrange multiplier for the incentive-compatibility constraint of agent \( i \). We can substitute for the resource constraint, and write the Lagrangean as:

\[
L(s_0, \omega, \omega^\infty, a^\infty, \lambda^\infty) = \\
= \sum_{i=1}^N \sum_{t=0}^\infty \beta^t \left\{ \phi_{it}(s^t) \left[ u(c_{it}(s^t)) - v(a_{it}(s^t)) \right] - \lambda_{it}(s^t) u'(a_{it}(s^t)) \right\} \Pi(s^t | s_0, a^{t-1}(s^{t-1}))
\]

where

\[
\phi_{i,t+1}(s^t, s_{t+1}) = \phi_{it}(s^t) + \lambda_{it}(s^t) \frac{\pi_{ai}(s_{i,t+1} | s_{it}, a_{it}(s^t))}{\pi(s_{i,t+1} | s_{it}, a_{it}(s^t))} \\
\phi_{it}(s_0) = \omega_i
\]

The new variables \( \phi_{it}(s^t) \), \( i = 1, ..., N \), are endogenously evolving Pareto-Negishi weights which have the same interpretation as in previous Sections. Mele (2008) characterizes in detail the optimal allocation, therefore I refer the reader to it for an economic analysis of this model.

It is possible to prove the recursivity of this problem in the space \((s_t, \phi_t)\) of endogenous Pareto weights and state of nature (see Appendix C for details). Due to the homogeneity properties of the value function, the relevant state space can be reduced:

\[
\frac{1}{\phi_1} J(s, \phi_1, ..., \phi_N) = J(1, \frac{\phi_2}{\phi_1}, ..., \frac{\phi_N}{\phi_1}) = \tilde{J}(1, \frac{\phi_2}{\phi_1}, ..., \frac{\phi_N}{\phi_1})
\]

therefore we only need \( N - 1 \) endogenous states.

In Section 5 I show a numerical example for the case of 2 agents, which is examined in detail by Zhao (2007) with APS techniques. With respect to the latter, my approach is much simpler, since the endogenous states are summarized by the ratio of the Pareto weights \( \theta \equiv \frac{\phi_N}{\phi_1} \) (Appendix C shows how to rewrite Lagrangean first-order conditions in terms of the ratio \( \theta \)). Therefore, in terms of numerical complexity, solving this model with my approach has the same difficulty of a stochastic neoclassical growth model.
4 Examples of models that are untractable under APS approach

I provide three examples of interesting economic problems that are untractable with promised utilities. One example is an extension of Abraham and Pavoni (forthcoming) to include human capital, and it can be interpreted as a model of optimal unemployment insurance in presence of non-monitorable financial market and human capital accumulation. The second is an example from the literature on executive compensation. Finally, I present a stylized model of international risk sharing with many countries and capital accumulation.

4.1 Unemployment insurance with human capital

The general result of Hopenhayn and Nicolini (1997) of a decreasing profile across time for unemployment benefits is well known. They assume there is no human capital accumulation in their setup, but it is a well documented fact that human capital depreciates during unemployment spells while increases during employment tenure. What happens to UI profile if we add human capital to Hopenhayn and Nicolini (1997)? Pavoni (forthcoming) shows that, in a model of UI with human capital depreciation, the general result of Hopenhayn and Nicolini (1997) of a decreasing profile for unemployment benefits survives, but there is a point at which benefits stop to decrease and are kept constant by the principal. This is due to depreciation of human capital during unemployment spells: for low levels of human capital, the principal has no interest in inducing the worker to find a new job.

Unemployment insurance models are kept simple for tractability reasons, but if we want a proper quantitative assessment of the welfare consequences of different schemes, we should have a more detailed framework. For example, we would like to analize the framework in Pavoni (forthcoming) when the agent can accumulate hidden assets as in the model in Section 3.1. While this extension is basically untractable with APS approach, my techniques can easily provide an answer.

To this purpose, take the model in Section 3.1 and assume the agent is endowed with initial human capital $h_{t-1}$. There are two possible state of the world: $S = \{U, E\}$ where $U$ means unemployed and $E$ means employed. Human capital is accumulated during employment spells by on-the-job training or learning-by-doing, while it depreciates during unemployment periods, according to the following process:

$$h_t (s^t) = \begin{cases} 
(1 + \rho) h_{t-1} (s^{t-1}) & s_t = E \\
(1 - \delta) h_{t-1} (s^{t-1}) & s_t = U 
\end{cases}$$

I assume that the wage is function of the human capital of the worker $y_t (s_t) = F (h_{t-1} (s^{t-1}))$. I also maintain the assumption in Pavoni (forthcoming) about the dependence of transition probabilities on the human capital level of the worker: $\pi (s_{t+1} | s_t, h_{t-1} (s^{t-1}), a_t (s^t))$. In this setup, the agent exerts effort when employed to increase the probability to keep the job; when unemployed, effort increases the probability
of finding a job, similar to Wang and Williamson (1996). The agent’s problem is then:

\[
V (s_0, h_{-1}, b_{-1}; \tau^\infty) = \max_{\{c_t(s^t), l_t(s^t), a_t(s^t), b_t(s^t)\}_{t=0}^\infty} \left\{ \sum_{t=0}^\infty \sum_{s^t} \beta^t [u (c_t (s^t)) - v (a_t (s^t))] \Pi (s^t | s_0, h^{t-2} (s^{t-2}) , a^{t-1} (s^{t-1})) \right\} \\
\text{s.t.} \quad c_t (s^t) + b_t (s^t) = E (h_{t-1} (s^{t-1})) + \tau_t (s^t) + Rh_{t-1} (s^{t-1})
\]

\[
h_t (s^t) = \begin{cases} 
(1 + \rho) h_{t-1} (s^{t-1}) & s_t = E \\
(1 - \delta) h_{t-1} (s^{t-1}) & s_t = U 
\end{cases}
\]

\[
h_t (s^t) \geq 0 \quad h_{-1} \text{ given}
\]

We interpret \( \tau_t (s^t) \) in the following way: when the worker is unemployed, \( \tau_t (s^t) \) is the unemployment benefit; when the worker is employed, \( \tau_t (s^t) \) is a tax. Accordingly, agent’s first order conditions with respect to the effort are

\[
v' (a_t (s^t)) = \beta R \sum_{s^{t+j} | s^t} \pi_j (s^{t+j} | s^t) \Pi (s^{t+j} | s^t, h^{t+j-1} (s^{t+j-1}) , a^{t+j} (s^{t+j} | s^t))
\]

and the first order condition with respect to bonds is:

\[
u' (a_t (s^t)) = \beta R \sum_{s^{t+j} | s^t} u' (c_{t+1} (s^{t+1})) \pi (s_{t+1} | s_t, h_{t-1} (s^{t-1}), a_t (s^t))
\]

Let \( \beta' \lambda_t (s^t) \) the Lagrange multiplier for (11), and \( \beta' \eta_t (s^t) \) the Lagrange multiplier for (12). I obtain the Lagrangean:

\[
L (s_0, \gamma, c^\infty, a^\infty, h^\infty, \lambda^\infty) = \sum_{t=0}^\infty \sum_{s^t} \beta^t \{ u (s_t) - c_t (s^t) + \phi_t (s^t) [u (c_t (s^t), l_t (s^t)) - v (a_t (s^t))] + \\
- \lambda_t (s^t) v' (a_t (s^t)) + [\eta_t (s^t) - \beta^{-1} \zeta_t (s^t)] u_c (c_t (s^t)) \} \Pi (s^t | s_0, h^{t-2} (s^{t-2}) , a^{t-1} (s^{t-1}))
\]

where

\[
\phi_{t+1} (s^t, \hat{s}) = \phi_t (s^t) + \lambda_t (s^t) \frac{\pi_a (s_{t+1} = \hat{s} | s_t, h_{t-1} (s^{t-1}) , a_t (s^t))}{\pi (s_{t+1} = \hat{s} | s_t, h_{t-1} (s^{t-1}) , a_t (s^t))} \quad \forall \hat{s} \in S \quad \text{and} \quad \phi_0 (s^0) = \gamma
\]

\[
\zeta_{t+1} (s^t, \hat{s}) = \eta_t (s^t) \quad \forall \hat{s} \in S \quad \text{and} \quad \zeta_0 (s^0) = 0
\]

\[
h_t (s^t) = \begin{cases} 
(1 + \rho) h_{t-1} (s^{t-1}) & s_t = E \\
(1 - \delta) h_{t-1} (s^{t-1}) & s_t = U 
\end{cases}
\]

\[17\]
Using the same arguments of Proposition 1, it is possible to show that the problem is recursively characterized by the Pareto weight $\phi_t (s')$, human capital $h_t (s')$ and costate variable $\zeta_t (s')$ (a proof is available upon request).

### 4.2 CEO compensation

There is a growing literature on executive compensation schemes, which builds on the general principal-agent model. For example, Clementi et al (2006) shows in a two-period model how stock compensation to executives can be thought as a commitment device for the firm: since the firm cannot credibly commit to pay a severance package, then stock grants are used to give deferred compensation to the manager (because it is more difficult to default on stockholders’ rights). Clementi et al. (2008a, 2008b) build a fully dynamic model with capital, and analyze some interesting properties of the model in line with data, but they do not include stocks as part of the compensation and they do not look at the no commitment case. The main difficulty in analyzing more realistic models is technical: the APS technique imposes a limit on the number of state variables that can be included in the setup.

In the following example, I describe a model of executive compensation with base salary, dividends and stocks for a firm that produces a good with capital, and where the manager has the option to leave each period. A continuum of risk-neutral investors own a firm which produces a good using capital

$$y_t (s') = A (s_t) f \left( k_{t-1} (s^{t-1}) \right)$$

where $A (s_t)$ is a productivity shock. The risk-averse manager of this firm can be compensated with base salary $w_t (s')$, and with stocks $\sigma_t (s') \leq 1$, where 1 is the total number of stocks. Manager’s effort is unobservable, and affects the probability distribution of productivity shock $A (s_t)$. The feasibility constraint for the firm is:

$$w_t (s') + d_t (s') + i_t (s') \leq A (s_t) f \left( k_{t-1} (s^{t-1}) \right), \quad k_{t-1} \text{ given}$$

where $i_t (s')$ is investment in physical capital and $d_t (s')$ are distributed dividends. The law of motion for capital is:

$$k_t (s') = i_t (s') + (1 - \delta) k_{t-1} (s^{t-1})$$

where $\delta$ is the depreciation rate of capital. Combining feasibility and the law of motion for capital, we get the following resource constraint:

$$w_t (s') + d_t (s') + k_t (s') - (1 - \delta) k_{t-1} (s^{t-1}) \leq A (s_t) f \left( k_{t-1} (s^{t-1}) \right)$$

I assume that the firm does not issue new stocks, and that manager’s consumption is perfectly monitorable. The latter assumption is equivalent to assume that the manager receives income only from the firm and cannot save or invest in other stocks or bonds. Therefore manager’s consumption $c_t (s')$ is given by:

$$c_t (s') = w_t (s') + \sigma_{t-1} (s^{t-1}) d_t (s') + p_t (s') (\sigma_t (s') - \sigma_{t-1} (s^{t-1}))$$

18
where $p_t (s')$ is the price of stocks. Since each investor is risk-neutral, and owns $1 - \sigma_t (s')$ stocks, then their budget constraint is

$$c_t' (s') - p_t (s') (\sigma_t (s') - \sigma_{t-1} (s^{t-1})) = d_t (s') (1 - \sigma_{t-1} (s^{t-1}))$$

and therefore, by imposing a no-bubble condition:

$$p_t (s') = \beta \sum_{s' \in S^{t+1}} [p_{t+1} (s'^{t+1}) + d_{t+1} (s'^{t+1})] \pi (s_{t+1} | s_t, a_t (s'))$$

The set of feasible contracts is then

$$\Gamma \equiv \left\{ (a^\infty, c^\infty, w^\infty, i^\infty, d^\infty, \sigma^\infty, k^\infty) : a_t (s') \in A, \quad c_t (s') \geq 0, \quad k_t (s') \in K \subseteq \mathbb{R}_+, \quad p_t (s') (\sigma_t (s') - \sigma_{t-1} (s^{t-1})) \geq 0, \quad \sigma_t (s') \in [0, 1], \quad c_t (s') + k_t (s') - (1 - \delta) k_{t-1} (s^{t-1}) \leq A (s_t) f (k_{t-1} (s^{t-1})) \quad \forall s' \in S^{t+1}, t \geq 0 \right\}$$

The problem of the manager is

$$\max_{\{c_t (s'), a_t (s')\}} \sum_{t=0}^{\infty} \beta^t \sum_{s'} [u (c_t (s')) - v (a_t (s'))] \Pi (s' | s_0, a^{t-1} (s^{t-1}))$$

s.t. $c_t (s') \leq w_3 (s') + \sigma_{t-1} (s^{t-1}) d_t (s') + p_t (s') (\sigma_t (s') - \sigma_{t-1} (s^{t-1}))$

From this maximization, we obtain the following first-order condition for effort:

$$v'_a (a_t (s')) = \sum_{t=1}^{\infty} \beta^t \sum_{s' \in S^{t+1}} \pi_{a_t} (s_{t+1} | s_t, a_t (s')) \times \left[ u (c_{t+j} (s'^{t+j})) - v (a_{t+j} (s'^{t+j})) \right] \Pi (s'^{t+j} | s', a'^{t+j-1} (s'^{t+j-1} | s'))$$

I assume that in each period the manager has the option to quit; if he quits, he immediately finds another job in another firm which is of the same size (i.e., has the same capital) of the first one. The investor sets a contract that makes quitting undesirable for the manager, i.e. I add a participation constraint to the
investor’s problem. The Pareto-constrained allocation can be found by solving:

\[
W(s_0, k_{-1}, \sigma_{-1}) = \max_{\{w_t(s^t), a_t(s^t), d_t(s^t), \sigma_t(s^t), k_t(s^t)\} \in \mathcal{C}_E \cap \mathcal{F}_E} \left\{ \sum_{t=0}^{\infty} \beta^t \sum_{s_t} [A(s_t) f(k_{t-1}(s^{t-1})) - w_t(s^t) -
- k_t(s^t) + (1 - \delta) k_{t-1}(s^{t-1}) - \sigma_{t-1}(s^{t-1}) d_t(s^t) -
- \mu_t(s^t)(\sigma_t(s^t) - \sigma_{t-1}(s^{t-1}))] \Pi(s^t | s_0, a_{-1}(s^{t-1})) \right\}
\]

s.t. \( v'_a(a_t(s^t)) = \sum_{j=1}^{\infty} \beta^j \sum_{s^{t+j} \mid s^t} \frac{\pi_a(s_{t+1} \mid s_t, a_t(s^t))}{\pi(s_{t+1} \mid s_t, a_t(s^t))} \times
\]
\[ \times \left[ u(c_{t+j}(s^{t+j})) - v(a_{t+j}(s^{t+j})) \right] \Pi(s^{t+j} \mid s^t, a^{t+j-1}(s^{t+j-1} \mid s^t)) \]

(13)

plus the following constraints

\[
p_t(s^t) = \sum_{j=1}^{\infty} \beta^j \sum_{s^{t+j} \mid s^t} \pi_a(s_{t+1} \mid s_t, a_t(s^t)) \times
\]
\[ \times \left[ u(c_{t+j}(s^{t+j})) - v(a_{t+j}(s^{t+j})) \right] \Pi(s^{t+j} \mid s^t, a^{t+j-1}(s^{t+j-1} \mid s^t)) \]

(14)

\[
w_t(s^t) + d_t(s^t) + \lambda_t(s^t) - (1 - \delta) k_{t-1}(s^{t-1}) \leq A(s_t) f(k_{t-1}(s^{t-1})), \quad k_{-1} \text{ given}
\]

\[
c_t(s^t) \leq w_t(s^t) + \sigma_{t-1}(s^{t-1}) d_t(s^t) + p_t(s^t)(\sigma_t(s^t) - \sigma_{t-1}(s^{t-1})), \quad \sigma_{-1} = 0
\]

(15)

\[
\sum_{j=0}^{\infty} \sum_{s^{t+j} \mid s^t} \beta^j \left[ u(c_{t+j}(s^{t+j})) - v(a_{t+j}(s^{t+j})) \right] \Pi(s^{t+j} \mid s^t, a^{t+j-1}(s^{t+j-1} \mid s^t)) \geq
\]
\[
V^{OUT}(s^t, k_{t-1}(s^{t-1}), 0) \quad \forall t \quad \forall s^t
\]

where \( V^{OUT}(s^t, k_{t-1}(s^{t-1}), 0) \) is the value for the manager of working in a new firm with the same capital as the one he leaves.

We associate a Lagrange multiplier \( \beta_t \lambda_t(s^t) \) to any ICC constraint (13), \( \beta_t \gamma_t(s^t) \) to the participation constraint (15), and \( \beta_t \mu_t(s^t) \) to the price constraint (14). We can now write down the Lagrangean:

\[
L(s_0, w^\infty_a, d^\infty_a, \sigma^\infty, k^\infty, \lambda^\infty, \gamma^\infty, \mu^\infty) = \sum_{t=0}^{\infty} \sum_{s^t} \beta^t \left\{ A(s_t) f(k_{t-1}(s^{t-1})) - w_t(s^t) - k_t(s^t) + (1 - \delta) k_{t-1}(s^{t-1}) +
- \sigma_{t-1}(s^{t-1}) d_t(s^t) - p_t(s^t)(\sigma_t(s^t) - \sigma_{t-1}(s^{t-1})) + \mu_t(s^t)p_t(s^t) - \gamma_t(s^t) \right\}
\]

(16)

\[
+ (\phi_t(s^t) + \gamma_t(s^t)) \left[ u(c_t(s^t)) - v(a_t(s^t)) \right] +
\]

\[
- \lambda_t(s^t)v'_a(a_t(s^t)) - \gamma_t(s^t) V^{OUT}(s^t, k_{t-1}(s^{t-1}), 0) \right\} \Pi(s^t | h_0, a_{-1}(s^{t-1}))
\]

20
where

\[
\phi_{t+1}(s', \hat{s}) = \phi_t(s') + \lambda_t(s') \frac{\pi_a(s_{t+1} = \hat{s} \mid s_t, a_t(s'))}{\pi(s_{t+1} = \hat{s} \mid s_t, a_t(s'))} + \gamma_t(s') \quad \forall \hat{s} \in S
\]

\[
\zeta_{t+1}(s', \hat{s}) = \zeta_t(s') + \mu_t(s') \quad \forall \hat{s} \in S
\]

\[
\zeta_0(s^0) = \phi_0(s^0) = 0
\]

As in the previous section, we can associate a saddle point functional equation to (a generalized version of) the Lagrangean, and show that the operator is a contraction (proof available upon request). In this case, therefore, the solution will be a Markovian policy function that depends on productivity shock, capital, stock holdings, the Pareto weight \(\phi_t(s')\) and the costate \(\zeta_t(s')\).

### 4.3 International risk sharing with moral hazard

Assume we live in a world with \(N\) countries. In each country, there is a representative entrepreneur endowed with an initial amount of capital \(k_{i,-1}\), \(i = 1, ..., N\) and with a production technology

\[
y_{it}(s') = A_i(s_i') f(k_{i,t-1}(s^{t-1}))
\]

where \(A_i(s_i')\) is a productivity shock and \(s_i'\) is the history of states of the world in country \(i\). Each entrepreneur enjoys consumption, and exerts unobservable effort that affects the productivity shock in the next period through the probability function \(\pi(s_{i,t+1} \mid s_{it}, a_{it}(s'))\). Assume the state of the world in each country is iid across time and across countries. Define \(s_t = (s_{1t}, s_{2t}, ..., s_{Nt})\). The budget constraint of entrepreneur of country \(i\) is:

\[
c_{it}(s') + i_{it}(s') + \sum_{j \neq i}^{N} n x_{it}^{ij}(s') \leq A_i(s') f(k_{i,t-1}(s^{t-1}))
\]

where \(i_{it}(s')\) is investment in physical capital, and \(n x_{it}^{ij}(s')\) is net exports of country \(i\) with country \(j\). The law of motion of capital is:

\[
k_{it}(s') = i_{it}(s') + (1 - \delta_i) k_{i,t-1}(s^{t-1})
\]

where \(\delta_i\) is the depreciation rate of capital in country \(i\). The world resource constraint becomes

\[
\sum_{i=1}^{N} c_{it}(s') + \sum_{i=1}^{N} k_{it}(s') - \sum_{i=1}^{N} (1 - \delta_i) k_{i,t-1}(s^{t-1}) \leq \sum_{i=1}^{N} A_i(s') f(k_{i,t-1}(s^{t-1}))
\]

\(k_{i,-1}\) given \(\forall i = 1, ..., N\)
In this Section, I present some numerical examples of the models described in previous Sections. The main advantage with respect to APS is that I do not need to characterize the feasible set for continuation values, which allows me to solve models with a large number of state variables. All the examples presented are not calibrated.

Notice that this model is an extension of the one presented in Section 3.2. The Pareto-constrained allocation can be found by solving:

$$P \left( s_0, \{ k_{i-1} \}_{i=1}^N \right) = \max_{\{ (c_{it}(s'), a_{it}(s'), k_{it}(s')) \}_{i=1}^N} \left\{ \sum_{i=1}^N \sum_{t=0}^\infty \beta^t \sum_{s'} \omega_i \left[ u(c_{it}(s')) - v(a_{it}(s')) \right] \Pi \left( s^t | s_0, a^{t-1} \left( s^{t-1} \right) \right) \right\}$$

s.t. $\nu' \left( a_{it}(s') \right) = \sum_{j=1}^\infty \sum_{s' \in \mathcal{S}} \beta^j \frac{\pi_{it} \left( s_{it+1} | s_{it}, a_{it}(s') \right)}{\pi \left( s_{it+1} | s_{it}, a_{it}(s') \right)} \left[ u(c_{it+j}(s'^{t+j})) - \nu \left( a_{it+j}(s'^{t+j}) \right) \right] \Pi \left( s^{t+j} | s', a^{t+j-1} \left( s^{t+j-1} \right) \right) \quad \forall i = 1, ..., N$

We can now write down the Lagrangean of this problem:

$$L(s^\infty, \nu^\infty, k^\infty, \phi^\infty) = \sum_{i=1}^N \sum_{t=0}^\infty \sum_{s'} \beta^t \left\{ \phi_{it} \left( s' \right) \left[ u(c_{it}(s')) - v(a_{it}(s')) \right] - \lambda_{it} \left( s' \right) \nu' \left( a_{it}(s') \right) \right\} \Pi \left( s^t | h_0, a^{t-1} \left( s^{t-1} \right) \right)$$

where

$$\phi_{it+1} \left( s', s_{it+1} \right) = \phi_{it} \left( s' \right) + \lambda_{it} \left( s' \right) \frac{\pi_{ai} \left( s_{it+1} | s_{it}, a_{it}(s') \right)}{\pi \left( s_{it+1} | s_{it}, a_{it}(s') \right)}$$

$$\phi_{it} \left( s_0 \right) = \omega_i \quad \forall i = 1, ..., N$$

The properties of this model are analyzed in Mele (2008). It can be easily shown that the Lagrangean is recursive in the $\left\{ \phi_{it} \left( s' \right) \right\}_{i=1}^N, \left\{ k_{i-1} \left( s^{t-1} \right) \right\}_{i=1}^N$. In the case of $N = 2$, it is possible to show that the state space can be reduced thanks to the homogeneity properties of the value and policy functions; when there are only two countries, then, the model is recursive in $\theta_1 \left( s' \right), k_{1-1} \left( s^{t-1} \right), k_{2-1} \left( s^{t-1} \right)$, where $\theta_1 \left( s' \right) \equiv \frac{\phi_2 \left( s' \right)}{\phi_1 \left( s' \right)}$

(All proofs are included in Mele (2008), or available upon request).

5 Numerical simulations: a new algorithm

In this Section, I present some numerical examples of the models described in previous Sections. The main advantage with respect to APS is that I do not need to characterize the feasible set for continuation values, which allows me to solve models with a large number of state variables. All the examples presented are not calibrated.
5.1 The algorithm

For simplicity, I assume that the Markov process has only two possible realizations \((s^L < s^H)\). I also assume there is no persistence across time (i.e., the state is i.i.d.), and I use the simpler notation \(\pi(a_t) = \pi(s_{t+1} = s^H \mid a_t)\). The numerical procedure is a collocation algorithm (see Judd (1998)) over the first-order conditions of the Lagrangean. From the recursive formulation we know that policy functions depend on the natural states of the problem and on the costates (i.e., Pareto weights) that come out from the Lagrangean approach. Let \(\varsigma\) be the vector of allocations, \(\chi\) be the vector of Lagrange multipliers, \(x \in X\) be the vector of natural states, and \(\theta \in \Theta\) be the vector of costates, and define \(R(s, \varsigma, \chi, x, \theta)\) as the objective function in the saddle point functional equation, and \(r(s, \varsigma, \chi, x, \theta)\) as the instantaneous utility function for the agent.

We therefore proceed as follows:

1. Fix \(\gamma\) and define a discrete grid \(G \subset X \times \Theta\) for natural states and costates.

2. Approximate policy functions for allocations \(\varsigma\) and Lagrange multipliers \(\chi\), the value function of the principal \(J\) and the continuation value of the agent \(U\) using cubic splines (or Chebychev polynomials, depending on the application), and set initial conditions for the approximation parameters\(^{11}\)

3. For any \((s, x, \theta) \in G\), use a nonlinear solver\(^{12}\) to solve for the Lagrangean first order conditions and the following two equations for the continuation value \(U\) and the value function \(J\):

\[
U(s, x, \theta) = r(s, \varsigma, \chi, x, \theta) + \beta \left[ \pi(a) U\left(s^H, x'^H, \theta'^H\right) + (1 - \pi(a)) U\left(s^L, x'^L, \theta'^L\right) \right] \\
J(s, x, \theta) = R(s, \varsigma, \chi, x, \theta) + \beta \left[ \pi(a) J\left(s^H, x'^H, \theta'^H\right) + (1 - \pi(a)) J\left(s^L, x'^L, \theta'^L\right) \right]
\]

\(^{17}\)\(^{18}\)

I use the Miranda-Fackler Compecon toolbox for function approximation; I check the degree of approximation by calculating the residuals of the Lagrangean first order conditions (i.e., how much the numerically approximated first-order conditions are different from zero) on a grid \(G^{test} \subseteq G\) with many gridpoints. In all applications, steps 1-3 are applied first to a grid with very few gridpoints, and then I increase the precision of the approximation by applying steps 1-3 to a finer grid. In general, a good approximation is obtained with few gridpoints. The algorithm is coded in Matlab.

5.1.1 Repeated moral hazard

In order to make the algorithm clear, I provide a detailed example of the procedure in the case of a standard repeated moral hazard setup. I simplify the notation by writing a generic variable as \(x_t\) instead of \(x_t(s_t)\).

I assume that the income process has two possible realizations \((y^L = y(s^L)\) and \(y^H = y(s^H))\). I also assume there is no persistence across time (i.e., the state is i.i.d.), and I use the simpler notation \(\pi(a_t) = \pi(y_{t+1} = y^H \mid a_t)\).

\(^{11}\)In the next subsection, I provide a clarification about why it is important to parametrize the value function and the continuation value, by means of an example.

\(^{12}\)In all applications presented in this paper, I use a version of the Broyden algorithm coded by Michael Reiter.
The Lagrangean becomes:

\[ L = \mathbb{E}_t^a \sum_{i=0}^{\infty} \beta^i \{ (y_{t+j} - c_{t+j}) + \phi_t \{ u_t (c_{t+j}) - v_t (a_{t+j}) \} \} - \lambda_t v'_t (a_t) \]

with

\[
\phi^H_{t+1} = \phi_t + \lambda_t \frac{\pi_a (a_t)}{\pi (a_t)}
\]

\[
\phi^L_{t+1} = \phi_t - \lambda_t \frac{\pi_a (a_t)}{1 - \pi (a_t)}
\]

\[ \phi_0 (s^0) = \gamma \]

where \( \mathbb{E}_t^a \) is the expectation operator over histories induced by the probability distribution \( \pi (a_t) \). The first-order conditions can be rewritten as

\[ c_t : \quad u'_t (c_t) = \frac{1}{\phi_t} \quad (19) \]

\[ a_t : \quad 0 = -\lambda_t v'' (a_t) - \phi_t v'_t (a_t) + \]

\[ + \pi_a (a_t) \beta E_{t+1}^u \left\{ \sum_{j=1}^{\infty} \beta^{i-1} \{ (y_{t+j} - c_{t+j}) - \lambda_{t+j} v'_t (a_{t+j}) + \phi_{t+j} [u_t (c_{t+j}) - v_t (a_{t+j})] \} | y_{t+1} = y^H \right\} + \]

\[ - \pi_a (a_t) \beta E_{t+1}^u \left\{ \sum_{j=1}^{\infty} \beta^{i-1} \{ (y_{t+j} - c_{t+j}) - \lambda_{t+j} v'_t (a_{t+j}) + \phi_{t+j} [u_t (c_{t+j}) - v_t (a_{t+j})] \} | y_{t+1} = y^L \right\} + \]

\[ + \beta \lambda_t \pi (a_t) \frac{\partial}{\partial a_t} \left[ \pi_a (a_t) - \pi (a_t) \right] [u_t (c_{t+1}) - v_t (a_{t+1}) | y_{t+1} = y^H] + \]

\[ + \beta \lambda_t (1 - \pi (a_t)) \frac{\partial}{\partial a_t} \left[ -\pi_a (a_t) (1 - \pi (a_t)) \right] [u_t (c_{t+1}) - v_t (a_{t+1}) | y_{t+1} = y^L] \]

and

\[ \lambda_t : \quad 0 = -v'_t (a_t) + \pi_a (a_t) \beta E_{t+1}^u \left\{ \sum_{j=1}^{\infty} \beta^{j-1} [u_t (c_{t+j}) - v_t (a_{t+j})] | y_{t+1} = y^H \right\} + \quad (21) \]

\[ - \pi_a (a_t) \beta E_{t+1}^u \left\{ \sum_{j=1}^{\infty} \beta^{j-1} [u_t (c_{t+j}) - v_t (a_{t+j})] | y_{t+1} = y^L \right\} \]

Notice that

\[ J (y^i, \phi^i_{t+1}) = \mathbb{E}_{t+1}^u \left\{ \sum_{j=1}^{\infty} \beta^{j-1} \{ (y_{t+j} - c_{t+j}) - \lambda_{t+j} v'_t (a_{t+j}) + \phi_{t+j} [u_t (c_{t+j}) - v_t (a_{t+j})] \} | y_{t+1} = y^i \right\} \]

\[ i = H, L \]
and

\[ U(y^i, \phi_{t+1}^i) = E_{t+1}^a \left\{ \sum_{j=1}^{\infty} \beta^{j-1} \left[ u(c_{t+j}) - v(a_{t+j}) | y_{t+1} = y^i \right] \right\} \]

\[ i = H, L \]

Therefore we can rewrite (20) and (21) as

\[ a_t : 0 = -\lambda_t u'(a_t) - \phi_t u'(a_t) + \beta \pi_a(a_t) \left[ J(y^H, \phi_{t+1}^H) - J(y^L, \phi_{t+1}^L) \right] + \]

\[ + \beta \lambda_t \left\{ \pi(a_t) \left[ \frac{\partial}{\partial a_t} \left( \frac{\pi_a(a_t)}{\pi(a_t)} \right) \right] \left[ u(c_{t+1}) - v(a_{t+1}) | y_{t+1} = y^H \right] + \right\} \]

\[ + (1 - \pi(a_t)) \left[ \frac{\partial}{\partial a_t} \left( \frac{-\pi_a(a_t)}{1-\pi(a_t)} \right) \left[ u(c_{t+1}) - v(a_{t+1}) | y_{t+1} = y^L \right] \right] \}

\[ \lambda_t : 0 = -v'(a_t) + \beta \pi_a(a_t) \left[ U(y^H, \phi_{t+1}^H) - U(y^L, \phi_{t+1}^L) \right] \]

(23)

I fix \( \gamma \) and I choose a discrete grid for \( \phi_t \) that contains \( \gamma \). I approximate with cubic splines \( a, \lambda, U \) and \( J \) on each grid node. I get consumption directly from \( \phi \) by using (19): \( c = u^{-1}(\phi^{-1}) \). Finally I solve the system of nonlinear equations that includes (22), (23), (17) and (18).

I choose the following functional forms:

\[ u(c) = \frac{c^{1-\sigma}}{1-\sigma} \]

\[ v(a) = \alpha a^\nu \]

\[ \pi(a) = a^\nu, \quad a \in (0,1) \]

The baseline parameters are summarized in the table:

<table>
<thead>
<tr>
<th>( \alpha )</th>
<th>( \varepsilon )</th>
<th>( \nu )</th>
<th>( \sigma )</th>
<th>( y^L )</th>
<th>( y^H )</th>
<th>( \beta )</th>
<th>( \gamma )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>2</td>
<td>0.5</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>0.95</td>
<td>0.5955</td>
</tr>
</tbody>
</table>

The algorithm delivers a set of parameterized policy functions. Figure 1 shows consumption, effort, the next period Pareto weights and the ICC Lagrange multiplier as functions of the current state \( \phi \). As we already said, consumption is increasing in \( \phi \), while effort is decreasing in the Pareto weight. Notice also that the policy functions for the Pareto weights satisfy Lemma 1. The Lagrange multiplier, interestingly, is an increasing function of the current state: as long as \( \phi \) increases (i.e., as long as the realizations of high income is preponderant), the shadow cost of enforcing an incentive compatible allocation decreases. Figure 2 represents the parameterized policy functions for transfers, continuation value of the agent and value function of the principal. Transfers are increasing in \( \phi \), as agent’s lifetime utility; at the contrary, planner value is monotone decreasing in the Pareto weight. Figure 3 and 4 show the average allocations across 50 thousands independent simulations for 200 periods, starting with \( y_0 = y^H \). In general, these simulations are in line
with previous studies\textsuperscript{13}: average consumption has a downward shape. Effort, on the other hand, increases on average. As in Thomas and Worrall (1990), the average path for agent’s lifetime utility is decreasing, while the Lagrange multiplier $\lambda$ is reduced on average along the optimal path. Interestingly, $\phi$ does not show a monotone pattern. To understand the last plot of Figure 4, let us notice that it is possible to derive the asset holdings of the principal from optimal allocations (Appendix D shows the details): according to the simulations, average assets must decrease across time. Finally, Figure 5 shows the Pareto frontier: it is decreasing and strictly concave.

5.1.2 An example with hidden assets

I maintain the same functional forms and parameters as in the previous example. Policy functions for consumption, agent lifetime utility and $\lambda$ depicted in Figure 6 and 7 are strictly increasing and concave in both costates, while effort is strictly decreasing and convex. The simulated series in Figure 8 and 9 confirm the results in Abraham and Pavoni: on average, consumption and lifetime utility increase across time, while effort decreases. Asset holdings (see Appendix D to see a description of the calculations needed to calculate them) also increase. Finally, Figure 10 shows the Pareto frontier for different $\zeta_0$ (the natural one is zero): it is decreasing and strictly concave. An application of the verification procedure described in the Appendix B shows that the first-order approach is also correct.

5.1.3 An example of risk sharing with moral hazard

I assume that there are two identical agents and they have the same weight in the social welfare function, and I maintain the same functional forms and parameters, except for income realizations:

<table>
<thead>
<tr>
<th>$\alpha_i$</th>
<th>$\varepsilon_i$</th>
<th>$\nu_i$</th>
<th>$\sigma_i$</th>
<th>$y_i^L$</th>
<th>$y_i^H$</th>
<th>$\beta$</th>
<th>$\omega_i$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>2</td>
<td>0.5</td>
<td>2</td>
<td>.4</td>
<td>.6</td>
<td>0.95</td>
<td>0.5</td>
</tr>
</tbody>
</table>

Also in this case, my results are in line with the theoretical and numerical findings of Zhao (2007) and Friedman (1998). Remember from Section 3.2 that the relevant state is the ratio of endogenous Pareto weights $\theta \equiv \frac{\phi_2}{\phi_1}$. Figures 11 and 12 show that agent 1 consumption and lifetime utility are decreasing in $\theta$ for any possible state of the world while effort is increasing in $\theta$, while the contrary happens to agent 2. Given that $\theta = \frac{u'(c_1)}{w'(c_2)}$ (see first-order conditions in Appendix C), we can use it as a measure of consumption inequality: Figure 13 and 14 show a sample path of 200 periods. Notice that $\theta$ is very persistent, confirming the theoretical result that $\theta$ evolves as a submartingale (see Zhao (2007), but also Mele (2008) for characterization of this property with the Lagrangean approach). Finally, Figure 15 shows a decreasing, strictly concave Pareto frontier.

\textsuperscript{13}The fact that each simulations starts with the same initial value for the shock explains the jump in period 1 for many series.
5.1.4 An example of international risk sharing (moral hazard and capital accumulation)

Also in this case, there are two identical agents and they have the same weight in the social welfare function. I use the same functional forms, and the following production function:

\[ f(k) = k^\gamma \]

and I change some parameters:

<table>
<thead>
<tr>
<th>( \alpha_i )</th>
<th>( \varepsilon_i )</th>
<th>( \nu_i )</th>
<th>( \sigma_i )</th>
<th>( A^L )</th>
<th>( A^H )</th>
<th>( \beta )</th>
<th>( \omega_i )</th>
<th>( \delta_i )</th>
<th>( \gamma_i )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.05</td>
<td>2</td>
<td>0.1</td>
<td>2</td>
<td>0.45</td>
<td>0.55</td>
<td>0.95</td>
<td>0.5</td>
<td>0.06</td>
<td>0.3</td>
</tr>
</tbody>
</table>

For brevity, I present only average over 50000 independent simulations. Figure 16 shows that risk sharing implies persistent inequality in welfare: there is one agent that puts more effort. This is confirmed in Figure 17, where \( \theta \) is increasing (on average) across time: consumption inequality increases. Notice that agents differ also in their capital holdings: there is one agent that accumulates more capital than the other. Also in this case, the Pareto frontier is strictly concave and decreasing (Figure 18).

6 Conclusions

I have presented a Lagrangean approach to repeated moral hazard problems, and an algorithm which is much faster than the traditional APS techniques. My methodology allows the researcher to deal with models with many states, and to calibrate the simulated series to real data in a reasonable amount of time. A huge class of models which are untractable under standard techniques can be easily addressed with my approach.

This method has many possible applications. Given the speed, the algorithm can be useful (as a time-saving technique) also for those models that are tractable with traditional techniques, but computationally burdensome. Dynamic agency problems with hidden effort and hidden assets are a good example: while we have a good qualitative idea of the main predictions of this model, to the best of my knowledge we still do not have a quantitative assessment in a calibrated economy. Mainly this is due to numerical difficulties. The Lagrangean approach offers a chance to overcome these limits: we can easily calibrate models and match data, in order to better understand various issues as e.g. consumption-saving anomalies, optimal unemployment insurance with assets accumulation, DSGE models with financial frictions.

However, the main gain of the Lagrangean method can be seen in more complicated setups, which are untractable with current state-of-the-art algorithms. Models of repeated moral hazard with heterogeneous agents and endogenous states are a good example: they require to solve the problem of each agent and aggregate the resulting individual optimal choices, then iterating until a general equilibrium is found. APS techniques are unmanageable even with just two endogenous states, while for my approach it would be a simple computational task.
Other issues in which the Lagrangean approach has an advantage in terms of complexity of the framework are optimal taxation theory in economies with private information, models of entrepreneurial choice, and models of banking and credit markets.

There is a price to pay, though: either restricting the class of models we analyze, by imposing some mild assumptions on primitives (e.e., Condition 1 and 2 in this work), or in some cases by verifying numerically the optimality of the solution. In any case, these seem small costs compared to the benefit of analyzing issues that are at present unmanageable.

References


Appendix A: Proofs

**Proposition 1** Fix an arbitrary constant $K > 0$ and let $K_\theta = \max \{K, K \|\theta\|\}$. The operator

$$
(T_K f)(s, \theta) = \min_{\chi > 0; \|\chi\| \leq K_\theta} \max_{a,c} \left\{ h(a, c, \theta, \chi, s) + \beta \sum_{s'} \pi(s' \mid s, a) f(s', \theta'(s')) \right\}
$$

s.t. $\theta'(s') = \theta + \chi \frac{\pi_a(s' \mid s, a)}{\pi(s' \mid s, a)} \forall s'$

is a contraction.

**Proof.** The space

$$
M = \{ f : S \times \mathbb{R}^2 \to \mathbb{R} \text{ s.t. } a) \quad \forall \alpha > 0 \quad f(\cdot, \alpha \theta) = \alpha f(\cdot, \theta)
$$

b) $f(s, \cdot)$ is continuous and bounded \}

will be our candidate, with norm

$$
\|f\| = \sup \{ |f(s, \theta)| : \|\theta\| \leq 1, s \in S \}
$$

Marcet and Marimon (2000) show that $M$ is a nonempty complete metric space. Now, fix a positive constant $K$ and let $K_\theta = \max \{K, K \|\theta\|\}$. Define the auxiliary operator

$$
(T_K f)(s, \theta) = \min_{\chi > 0; \|\chi\| \leq K_\theta} \max_{a,c} \left\{ h(a, c, \theta, \chi, s) + \beta \sum_{s'} \pi(s' \mid s, a) f(s', \theta'(s')) \right\}
$$

s.t. $\theta'(s') = \theta + \chi \frac{\pi_a(s' \mid s, a)}{\pi(s' \mid s, a)} \forall s'$

I have to show that $T_K : M \to M$. Notice that

$$
(T_K f)(s, \theta) = \theta h_0(a^*, c^*, s) + \chi^* h_1(a^*, c^*, s) + \beta \sum_{s'} \pi(s' \mid s, a^*) f(s', \theta'(s'))
$$

hence by Schwartz’s inequality

$$
\|(T_K f)(s, \theta)\| \leq \|\theta\| \|h_0(a^*, c^*, s)\| + \max \{K, K \|\theta\|\} \|h_1(a^*, c^*, s)\| + \beta \left( \max \{K, K \|\theta\|\} \left\| \frac{\pi_a(s' \mid s, a^*)}{\pi(s' \mid s, a^*)} \right\| + \|\theta\| \right) \left\| f(s', \theta'(s')) \right\|}
$$

and therefore $(T_K f)(s, \phi)$ is bounded. A generalized Maximum Principle argument gives continuity of $(T_K f)(s, \phi)$. To check for homogeneity properties, let $(a^*, c^*, \chi^*)$ be such that

$$(T_K f)(s, \theta) = h(a^*, c^*, \theta, \chi^*, s) + \beta \sum_{s'} \pi(s' \mid s, a^*) f(s', \theta'(s'))$$

Then for any $\alpha > 0$ we get

$$
\alpha (T_K f)(s, \theta) = \alpha \left[ h(a^*, c^*, \theta, \chi^*, s) + \beta \sum_{s'} \pi(s' \mid s, a^*) f(s', \theta'(s')) \right]
$$
Therefore
\[
\begin{align*}
h (a^*, c^*, a\theta, a\chi^*, s) + \beta \sum_{s'} \pi (s' | s, a^*) f \left( s', a\theta^* (s') \right) \\
= \alpha \left[ h (a^*, c^*, \theta, \chi^*, s) + \beta \sum_{s'} \pi (s' | s, a^*) f \left( s', \theta^* (s') \right) \right]
\end{align*}
\]

Now take \( \chi, \theta' (s') = \varphi (a\theta, \chi, s') \) and a feasible \( a \). We can write:
\[
\begin{align*}
h (a^*, c^*, a\theta, \chi, s) + \beta \sum_{s'} \pi (s' | s, a^*) f \left( s', \theta^* (s') \right) \\
= \alpha \left[ h (a^*, c^*, \theta, \chi, s) + \beta \sum_{s'} \pi (s' | s, a^*) f \left( s', \theta^* (s') \right) \right]
\end{align*}
\]

and therefore
\[
\begin{align*}
(T_K f)(s, a\theta) &= h (a^*, c^*, a\theta, a\chi^*, s) + \beta \sum_{s'} \pi (s' | s, a^*) f \left( s', a\theta^* (s') \right) \\
&= \alpha (T_K f)(s, \theta)
\end{align*}
\]

and therefore the operator preserves the homogeneity properties. To see monotonicity, let \( g, u \in M \) such that \( g \leq h \). Therefore
\[
\begin{align*}
\max_{a, c} \left\{ h (a, c, \theta, \chi, s) + \beta \sum_{s'} \pi (s' | s, a) g \left( s', \theta' (s') \right) \right\} \\
\leq \max_{a, c} \left\{ h (a, c, \theta, \chi, s) + \beta \sum_{s'} \pi (s' | s, a) u \left( s', \theta' (s') \right) \right\}
\end{align*}
\]

and then
\[
\begin{align*}
\min_{\{\chi \geq 0, ||\chi|| \leq K_\alpha\}} \max_{a, c} \left\{ h (a, c, \theta, \chi, s) + \beta \sum_{s'} \pi (s' | s, a) g \left( s', \theta' (s') \right) \right\} \\
\leq \min_{\{\chi \geq 0, ||\chi|| \leq K_\alpha\}} \max_{a, c} \left\{ h (a, c, \theta, \chi, s) + \beta \sum_{s'} \pi (s' | s, a) u \left( s', \theta' (s') \right) \right\}
\end{align*}
\]

which implies \((T_K g)(s, \theta) \leq (T_K u)(s, \theta)\). To see discounting, let \( k \in \mathbb{R}_+ \), and define \( f + k \in M \) as \((f + k)(s, \theta) = f(s, \theta) + k\). Therefore:
\[
\begin{align*}
\max_{a, c} \left\{ h (a, c, \theta, \chi, s) + \beta \sum_{s'} \pi (s' | s, a) (g + k) \left( s', \theta' (s') \right) \right\} \\
= \max_{a, c} \left\{ h (a, c, \theta, \chi, s) + \beta \sum_{s'} \pi (s' | s, a) g \left( s', \theta' (s') \right) + \beta k \right\}
\end{align*}
\]

\[
= \max_{a, c} \left\{ h (a, c, \theta, \chi, s) + \beta \sum_{s'} \pi (s' | s, a) g \left( s', \theta' (s') \right) + \beta k \right\}
\]

32
Hence we get

\[ T_K (f + k) (s, \theta) = \min_{\chi \geq 0 \| \chi \leq K} \max_{a,c} \left\{ \sum_{s'} h(a, c, \theta, s, s') + \beta \sum_{s'} \pi(s' \mid s, a) f(s', \theta') \right\} + \beta k \]

and then \( T_K (f + k) \leq T_K f + \beta k \). Now it is possible to use the above properties to show the contraction property for the operator \( T_K \). In order to see this, let \( f, g \in M \). By homogeneity, we get

\[ f(s, \phi) = g(s, \theta) + f(s, \theta) - g(s, \theta) \leq g(s, \theta) + \| f(s, \theta) - g(s, \theta) \| \]

and then

\[ f(s, \theta) \leq g(s, \theta) + \| f(s, \theta) - g(s, \theta) \| \]

Now applying the operator \( T_K \) and using monotonicity and discounting we get:

\[ (T_K f)(s, \theta) \leq T_K (g + \| f - g \|)(s, \theta) \leq (T_K g)(s, \theta) + \beta \| f - g \| \]

which implies finally

\[ \| T_K f - T_K g \| \leq \beta \| f - g \| \]

and given \( \beta \in (0, 1) \) this concludes the proof that the operator \( T_K \) is a contraction.

\[ \square \]

**Lemma 1** In the optimal contract, \( \phi_{t+1}(s', \tilde{s}_i) < \phi_t(s') < \phi_{t+1}(s', \tilde{s}_j) \) for any \( t \).

**Proof.** Notice first that, for any \( t \), \( \exists i, j : \pi_a(\tilde{s}_j \mid s_i, a_t(s')) > 0 \) and \( \pi_a(\tilde{s}_j \mid s_i, a_t(s')) < 0 \). Suppose not: then the only possibility is that \( \pi_a(\tilde{s}_i \mid s_i, a_t(s')) = 0 \) for any \( i \) (otherwise, \( \sum_{\tilde{s}_i} \pi_a(\tilde{s}_i \mid s_i, a_t(s')) \neq 0 \), which is impossible). This implies, by (7), \( 0 = v'(a_t(s')) \) which is a contradiction since \( v(\cdot) \) is strictly increasing. Adding the full support assumption and the fact that \( \lambda_t(s') > 0 \), we get that \( \exists i, j : \phi_{t+1}(s', \tilde{s}_j) < \phi_t(s') < \phi_{t+1}(s', \tilde{s}_i) \). By MLRC, \( \phi_{t+1}(s', \tilde{s}_1) \leq \phi_{t+1}(s', \tilde{s}_i) \) for any \( j \) and \( \phi_{t+1}(s', \tilde{s}_i) \leq \phi_{t+1}(s', \tilde{s}_j) \) for any \( i \), which proves the statement. \( \square \)

**Proposition 2** \( \phi_t(s') \) is a martingale that converges to zero.

**Proof.** Use the law of motion of \( \phi_t(s') \) and take expectations on both sides:

\[ \sum_{s_{t+1}} \phi_{t+1}(s', s_{t+1}) \pi(s_{t+1} \mid s_t, a_t(s')) = \phi_t(s') + \lambda_t(s') \sum_{s_{t+1}} \pi_a(s_{t+1} \mid s_t, a_t(s')) \pi(s_{t+1} \mid s_t, a_t(s')) \]
Notice that \( \lambda_t (s') \sum_{s_{t+1}} \frac{\pi_t(s_{t+1} | s_t, a_t(s'))}{\pi_t(s_{t+1} | s_t, a_t(s'))} \pi_t(s_{t+1} | s_t, a_t(s')) = 0 \), which implies

\[
E_t^a [\phi_{t+1} | s^t] = \phi_t (s')
\]

where \( E_t^a [\cdot] \) is the expectation operator induced by \( a_t (s') \). Therefore \( \phi_t (s') \) is a martingale. To see that it converges to zero, rewrite (24) by using (6):

\[
E_t^a \left[ \frac{1}{u_c (c_{t+1} (s_{t+1}))} \right] = \frac{1}{u_c (c_t (s'))}
\]

By Inada conditions, \( \frac{1}{u_c (c_t (s'))} \) is bounded above zero and below infinity. Therefore \( \phi_t (s') \) is a nonnegative martingale, and by Doob’s theorem it converges almost surely to a random variable (call it \( X \)). To see that \( X = 0 \) almost surely, I follow the proof strategy of Thomas and Worrall (1990), to which I refer for details. Suppose not, and take a path \( \{ s^t \}_{t=0}^\infty \) such that \( \lim_{t \to \infty} \phi_t (s^t) = \phi > 0 \) and state \( \hat{s}_I \) happens infinitely many times. I claim that this sequence cannot exist. Take a subsequence \( \{ s^t(k) \}_{k=1}^\infty \) of \( \{ s^t \}_{t=0}^\infty \) such that \( s^t(k) = \hat{s}_I \forall k \). This subsequence has to converge to some limit \( \overline{s} \) for some \( \epsilon > 0 \). Call \( f (\phi_t (s'), \hat{s}_I) = \phi_{t+1} (s', \hat{s}_I) \) and notice that \( f (\cdot) \) is continuous, hence \( \lim_{k \to \infty} f (\phi_t (s^t(k)), \hat{s}_I) = f (\overline{s}, \hat{s}_I) \). By definition, \( f (\phi_t (s^t(k)), \hat{s}_I) = \phi_{t+1} (s^t(k), \hat{s}_I) \), then \( \lim_{k \to \infty} \phi_{t+1} (s^t(k), \hat{s}_I) = f (\overline{s}, \hat{s}_I) \). However, notice that it must be \( \lim_{k \to \infty} \phi_{t(k)} (s^{t(k)}) = \overline{s} \) and \( \lim_{k \to \infty} \phi_{t(k)+1} (s^{t(k)}, \hat{s}_I) = \overline{s} \). But by Lemma 1, \( \phi_{t(k)} (s^{t(k)}) < \phi_{t(k)+1} (s^{t(k)}, \hat{s}_I) \) for any \( k \). Therefore, this is a contradiction and this sequence cannot exist. Since paths where state \( \hat{s}_I \) occurs only a finite number of times have probability zero, this implies that

\[
\Pr \left\{ \lim_{t \to \infty} \phi_t (s^t) > 0 \right\} = 0
\]

which implies \( X = 0 \) almost surely. \( \blacksquare \)
Appendix B: Model with hidden assets

6.1 Recursivity

Define the following generalized version of the problem:

\[
W_{\theta}^{SWF}(s_0) = \max_{\{a_t(s_t), e_t(s_t')\}_{t=0}^{\infty} \in \Gamma^{HA}} \sum_{t=0}^{\infty} \sum_{s_t} \beta^t \left[ g(s_t) - c_t(s_t) \right] \Pi(s_t | s_0, a_{t-1}(s_{t-1})) + \\
\gamma \sum_{t=0}^{\infty} \sum_{s_t} \beta^t \left[ u(c_t(s_t)) - v(a_t(s_t)) \right] \Pi(s_t | s_0, a_{t-1}(s_{t-1}))
\]

s.t. \( u'(a_t(s_t)) = \sum_{j=1}^{\infty} \beta^j \sum_{s_{t+j} \in s_t} \frac{\pi_a(s_{t+j} | s_t, a_t(s_t))}{\pi(s_{t+j+1} | s_t, a_t(s_t))} \times \left[ u(c_{t+j}(s_{t+j})) - v(a_{t+j}(s_{t+j})) \right] \Pi(s_{t+j} | s_t, a_{t+j-1}(s_{t+j-1} | s_t)) \)

\[ \forall s_t, t \geq 0 \]

\[ u'(c_t(s_t)) = \beta R \sum_{s_{t+1}} u'(c_{t+1}(s_t, s_{t+1})) \pi(s_{t+1} | s_t, a_t(s_t)) \]

The Lagrange is:

\[
L_\theta(s_0, \gamma, c_0^\infty, a_0^\infty, \lambda^\infty, \eta^\infty) = \sum_{t=0}^{\infty} \sum_{s_t} \beta^t \left[ \phi^t \left[ g(s_t) - c_t(s_t) \right] + \gamma \left[ u(c_t(s_t)) - v(a_t(s_t)) \right] \right] \Pi(s_t | s_0, a_{t-1}(s_{t-1})) + \\
- \sum_{t=0}^{\infty} \sum_{s_t} \beta^t \lambda_t(s_t) \left[ u'(a_t(s_t)) - \sum_{j=1}^{\infty} \beta^j \sum_{s_{t+j} \in s_t} \frac{\pi_a(s_{t+j} | s_t, a_t(s_t))}{\pi(s_{t+j+1} | s_t, a_t(s_t))} \times \left[ u(c_{t+j}(s_{t+j})) - v(a_{t+j}(s_{t+j})) \right] \Pi(s_{t+j} | s_t, a_{t+j-1}(s_{t+j-1} | s_t)) \right] \Pi(s_t | s_0, a_{t-1}(s_{t-1})) + \\
+ \sum_{t=0}^{\infty} \sum_{s_t} \beta^t \eta_t(s_t) \left[ u'(c_t(s_t)) - \sum_{s_{t+1}} u'(c_{t+1}(s_t, s_{t+1})) \pi(s_{t+1} | s_t, a_t(s_t)) \right] \Pi(s_t | s_0, a_{t-1}(s_{t-1}))
\]

Notice that \( r(a, c, s) \equiv g(s) - c \) is uniformly bounded by debt limits, therefore there exists a lower bound \( \kappa \) such that \( r(a, c, s) \geq \kappa \). As before, we can define \( \kappa < \frac{1}{\beta}, \varphi^1(\phi, \lambda, s') \equiv \phi + \lambda \frac{\pi_a(s' | s, a)}{\pi(s' | s, a)}, \varphi^2(\zeta, \eta, s') \equiv \eta, \)

\[
\Psi(\phi, \zeta, \eta, s') \equiv \begin{cases} 
\varphi^1(\phi, \lambda, s') \\
\varphi^2(\zeta, \eta, s')
\end{cases}, \quad h_0^P(a, c, s) \equiv r(a, c, s), \quad h_0^E(a, c, s) \equiv r(a, c, s) - \kappa, \quad h_0^{ICC}(a, c, s) \equiv u(c) - v(a), \quad h_1^{ICC}(a, c, s) \equiv -v'(a), \quad h_1^E(a, c, s) \equiv -\beta^{-1}u'_c(c), \quad h_1^{EE}(a, c, s) \equiv u'_c(c), \quad \theta \equiv \begin{bmatrix} \phi^0 & \phi & \zeta \end{bmatrix} \in \mathbb{R}^3,
\]

\[ \chi \equiv \begin{bmatrix} \lambda^0 & \lambda & \eta \end{bmatrix} \text{ and } \]

\[ h(a, c, \theta, \chi, s) \equiv \theta h_0(a, c, s) + \chi h_1(a, c, s) \]

\[ \equiv \begin{bmatrix} \phi^0 & \phi & \zeta \end{bmatrix} \begin{bmatrix} h_0^P(a, c, s) \\
 h_0^{ICC}(a, c, s) \\
 h_0^E(a, c, s) \end{bmatrix} + \begin{bmatrix} \lambda^0 & \lambda & \eta \end{bmatrix} \begin{bmatrix} h_1^P(a, c, s) \\
 h_1^{ICC}(a, c, s) \\
 h_1^E(a, c, s) \end{bmatrix} \]

35
which is homogenous of degree 1 in \((\theta, \chi)\). The Lagrangean can be written as:

\[
L_{\theta}(s_0, \gamma, \chi, \alpha, \lambda, \eta) = \sum_{t=0}^{\infty} \sum_{s_t} \beta_t h(a_t(s_t), c_t(s_t), \theta_t(s_t), \chi_t(s_t), s_t) \Pi(s_t \mid s_0, a_t^{-1}(s_t-1))
\]

where

\[
\theta_{t+1}(s', \tilde{s}) = \Psi(\theta_t(s'), \chi_t(s'), \tilde{s}) \quad \forall \tilde{s} \in S
\]

\[
\theta_0(s^0) = \begin{bmatrix} \theta^0 & \gamma \end{bmatrix}
\]

We can associate a saddle point functional equation to this Lagrangean

\[
J(s, \theta) = \min_{\chi} \max_{a,c} \left\{ h(a,c,\theta,\chi,s) + \beta \sum_{s'} \pi(s' \mid s,a) J(s', \theta'(s')) \right\}
\]

s.t. \( \theta'(s') = \Psi(\theta, \chi, s') \quad \forall s' \)

The following Proposition shows that the RHS operator is a contraction mapping.

**Proposition 3** Fix an arbitrary constant \( K > 0 \) and let \( K_{\theta} = \max \{ K, \|\theta\| \} \). The operator

\[
(T_Kf)(s, \theta) = \min_{\chi : \|\chi\| \leq K_{\theta}} \max_{a,c} \left\{ h(a,c,\theta,\chi,s) + \beta \sum_{s'} \pi(s' \mid s,a) f(s', \theta'(s')) \right\}
\]

s.t. \( \theta'(s') = \Psi(\theta, \chi, s') \quad \forall s' \)

is a contraction.

**Proof.** Straightforward by repeating the steps to prove Proposition 1 in the following space of functions:

\[
M = \{ f : S \times \mathbb{R}^3 \to \mathbb{R} \text{ s.t.} \}
\]

\[
a) \quad \forall \alpha > 0 \quad f(\cdot, \alpha \theta) = \alpha f(\cdot, \theta) \\
b) \quad f(s, \cdot) \text{ is continuous and bounded} \}
\]

with norm

\[
\|f\| = \sup \{|f(s, \theta)| : \|\theta\| \leq 1, s \in S\}
\]

6.2 First-order conditions for the hidden asset model

\[
L(s_0, \gamma, \chi, \alpha, \lambda, \eta) = \sum_{t=0}^{\infty} \sum_{s_t} \beta_t \left\{ y(s_t) - c_t(s_t) + \phi_t(s_t) \left[v(c_t(s')) - v(a_t(s'))\right]\right. \\
- \lambda_t(s') v'(a_t(s')) + \left[\eta_t(s') - \beta^{-1} \zeta_t(s')\right] u_c(c_t(s'))} \Pi(s_t \mid s_0, a_t^{-1}(s_t-1))
\]
\[ \alpha_t(s^t) : 0 = -1 + \phi_t(s^t) u_c(c_t(s^t)) + [\eta_t(s^t) - \beta^{-1} \zeta_t(s^t)] u_{cc}(c_t(s^t)) \]

\[ a_t(s^t) : 0 = -\lambda_t(s^t) v''(a_t(s^t)) - \phi_t(s^t) v'(a_t(s^t)) + \sum_{j=1}^{\infty} \beta^j \sum_{s^{t+j} \mid s^t} \frac{\pi_t(s_{t+1} \mid s_t, a_t(s^t))}{\pi_t(s_{t+1} \mid s_t, a_t(s^t))} \{ y(s_t) - c_t(s^t) - \lambda_{t+j}(s_{t+j}) v'(a_{t+j}(s^{t+j})) \}
+ \phi_{t+j}(s^{t+j}) [u(c_{t+j}(s^{t+j})) - v(a_{t+j}(s^{t+j}))] + [\eta_{t+j}(s^{t+j}) - \beta^{-1} \zeta_{t+j}(s^{t+j})] u_{cc}(c_{t+j}(s^{t+j})) \}
\times \pi_t(s_{t+1} \mid s_t, a_t(s^t)) \]

and

\[ \lambda_t(s^t) : 0 = -v'(a_t(s^t)) + \sum_{j=1}^{\infty} \sum_{s^{t+j} \mid s^t} \frac{\pi_t(s_{t+1} \mid s_t, a_t(s^t))}{\pi_t(s_{t+1} \mid s_t, a_t(s^t))} \times \left[ \beta^j [u(c_{t+j}(s^{t+j})) - v(a_{t+j}(s^{t+j}))] \right] \pi_t(s_{t+1} \mid s_t, a_t(s^t)) \]

\[ \eta_t(s^t) : 0 = u'(c_t(s^t)) - \sum_{s^{t+1} \mid s^t} u'(c_{t+1}(s^t, s_{t+1})) \pi_t(s_{t+1} \mid s_t, a_t(s^t)) \]

6.3 The verification procedure

There are not known conditions under which the first-order approach is valid in the framework with hidden effort and hidden assets. Therefore, we cannot be sure that a first-order approach delivers the correct optimal allocation; it is possible that the solution obtained does not satisfies the true incentive compatibility constraint of the original problem. However we can verify it by a simple numerical procedure similar to the one proposed by Abraham and Pavoni (forthcoming): we remaximize the lifetime utility of the agent, by taking as given the optimal transfer scheme implied by the solution of the Pareto problem; if remaximization delivers a welfare gain to the agent, the solution obtained with first-order approach does not satisfy incentive compatibility. Instead, if no gain is possible, then the first-order approach is valid.
We solve the following problem:

\[ V(s_0, b_{-1}, \gamma, 0) = \max_{\{c^V_t(s^t), a^V_t(s^t), b^V_t(s^t)\}_{t=0}^{\infty}} \left\{ \sum_{t=0}^{\infty} \sum_{s^t} \beta^t [u(c^V_t(s^t)) - v(a^V_t(s^t))] \Pi(s^t | s_0, a^{V_{t-1}}(s^{t-1})) \right\} \]

subject to

\[ c^V_t(s^t) + b^V_t(s^t) = y(s_t) + T(s_t, \phi_t(s^t), \zeta_t(s^t)) + Rb^V_{t-1}(s^{t-1}) \quad \text{and} \quad b_{-1} \text{ given} \]

\[ \phi_{t+1}(s^t, \bar{s}) = \varphi^1(\bar{s}, \phi_t(s^t), \zeta_t(s^t)) \quad \forall \bar{s} \in S \quad \text{and} \quad \phi_0(s^0) = \gamma \]

\[ \zeta_{t+1}(s^t, \bar{s}) = \varphi^2(\bar{s}, \phi_t(s^t), \zeta_t(s^t)) \quad \forall \bar{s} \in S \quad \text{and} \quad \zeta_0(s^0) = 0 \]

where \(T(\cdot), \varphi^1(\cdot)\) and \(\varphi^2(\cdot)\) are the policy functions derived from Lagrangean (9), and are exogenous from the point of view of the agent (they define the transfer policy of the principal). It is obvious that this problem is recursive in the state space \((s, \phi, \zeta, b)\), but notice that \(\phi\) and \(\zeta\) are exogenous states. Since the problem of the agent is strictly concave, first-order conditions are necessary and sufficient to get the policy functions.

We can get the optimal strategies by solving for

\[ u'(c^V_t(s^t)) = \beta R \sum_{s_{t+1}} u'(c^V_{t+1}(s^t, s_{t+1})) \pi(s_{t+1} | s_t, a^V_t(s^t)) \]

\[ v'(a^V_t(s^t)) = \sum_{j=1}^{\infty} \beta^j \sum_{s^{t+j} | s^t} \pi_a(s_{t+1} | s_t, a^V_t(s^t)) \times \]

\[ \times [u(c^V_{t+j}(s^{t+j})) - v(a^V_{t+j}(s^{t+j}))] \Pi(s^{t+j} | s^t, a^{V_{t+j-1}}(s^{t+j-1} | s^t)) \]

plus the budget constraint and the law of motion \(\varphi^1(\cdot)\) and \(\varphi^2(\cdot)\) for the costate variables. Once we get the policy functions and the value function, we can calculate the welfare gain from reoptimization with respect to the optimal allocation obtained with first-order approach: if this difference is zero (in numerical terms), then the Lagrangean first-order method delivers the solution of the original problem.
Appendix C: Model with multiple agents

6.4 Recursivity

First of all, notice that our problem is already in the form of a SWF maximizations, therefore we can apply directly the saddle-point functional equation to it. In this case, let
\[ r(a_i, c_i, s) \equiv u(c_i) - v(a_i), \]
\[ \varphi(\phi_i, \lambda_i, s') \equiv \phi_i + \lambda_i \frac{a_i(\phi_i|s_i, a_i)}{\pi(a_i|s_i, a_i)}, \]
\[ h_0(a, c, s) \equiv r(a_i, c_i, s), \]
\[ h_1(a, c, s) \equiv -v'(a_i), \]
\[ h(a, c, \phi, \lambda, s) \equiv \phi h_0(a, c, s) + \lambda h_1(a, c, s) \]
which is homogenous of degree 1 in \((\phi, \lambda)\). The Lagrangean can be written as:
\[
L(s, \omega, c, a, \lambda) = \sum_{t=0}^{\infty} \sum_{s_t} \beta_t h(a_t(s_t), c_t(s_t), \phi_t(s_t), \lambda_t(s_t), s_t) \prod(s_t | s_0, a_{t-1}(s_{t-1}))
\]
where
\[
\varphi_t+1(s_t, \tilde{s}) = \varphi(\phi_t(s_t), \lambda_t(s_t), \tilde{s}) \quad \forall \tilde{s} \in S
\]
\[
\phi_0(s^0) = \omega
\]
where of course \(\phi \in \mathbb{R}^N\). We can associate a saddle point functional equation to this Lagrangean
\[
J(s, \phi) = \min_{\chi} \max_{a, c} \left\{ h(a, c, \phi, \chi, s) + \beta \sum_{s'} \pi(s' | s, a) J(s', \phi'(s')) \right\} \quad \text{(SPFE)}
\]
s.t. \(\phi'(s') = \phi + \lambda \pi(a(s' | s, a)) \prod(s' | s, a) \forall s'
\]
Again, we are left to prove the following:

**Proposition 4** Fix an arbitrary constant \(K > 0\) and let \(K_\theta = \max \{K, K \|\phi\|\}\). The operator
\[
(T_K f)(s, \phi) = \min_{\lambda > 0} \max_{\|\phi\| \leq K_\theta} \left\{ h(a, c, \phi, \lambda, s) + \beta \sum_{s'} \pi(s' | s, a) f(s', \phi'(s')) \right\}
\]
s.t. \(\phi'(s') = \phi + \frac{a(s' | s, a)}{\pi(s' | s, a)} \forall s'
\]
is a contraction.

**Proof.** Straightforward by repeating the steps to prove Proposition 1 in the following space of functions:
\[
M = \{ f : S \times \mathbb{R}^N \rightarrow \mathbb{R} \quad \text{s.t.} \}
\]
a) \(\forall \alpha > 0 \quad f(\cdot, \alpha \phi) = \alpha f(\cdot, \phi)\)
b) \(f(s, \cdot) \text{ is continuous and bounded} \}
\]
with norm
\[
\|f\| = \sup \{ |f(s, \phi)| : \|\phi\| \leq 1, s \in S \}
\]

The next Subsection shows first-order conditions for the case where \(N = 2\).
6.5 First-order conditions (N=2)

The Lagrangean is

\[
\begin{align*}
L (s_0, \omega, c^\infty, a^\infty, \lambda^\infty) &= \\
&= \sum_{i=1}^{2} \sum_{t=0}^{\infty} \sum_{s^t} \beta^t \{ \phi_{it} (s^t) [u (c_{it} (s^t)) - u (a_{it} (s^t))] \} - \\
&- \lambda_{it} (s^t) u' (a_{it} (s^t)) \Pi (s^t | h_0, a^{t-1} (s^{t-1})) \\
\end{align*}
\]

\[
\begin{align*}
/ c_{it} (s^t) : & \quad \phi_{it} (s^t) u' (c_{it} (s^t)) = \phi_{2it} (s^t) u' (c_{2it} (s^t)) \\
/ a_{it} (s^t) : & \quad \phi_{it} (s^t) u' (a_{it} (s^t)) + \lambda_{it} (s^t) u'' (a_{it} (s^t)) = \sum_{i=1}^{2} \sum_{j=1}^{\infty} \sum_{s^{t+j} | s^t} \beta^j \{ \phi_{i,t+j} (s^t) [u (c_{i,t+j} (s^{t+j})) - \\
&- v (a_{i,t+j} (s^{t+j}))] - \lambda_{i,t+j} (s^{t+j}) u' (a_{i,t+j} (s^{t+j})) \} \times \\
&\times \pi_{ai} (s_{i,t+1} | s_{it}, a_{it} (s^t)) \Pi (s^{t+j} | s_{t}, a^{t+j-1} (s^{t+j-1})) + \\
&+ \beta \lambda_{it} (s^t) \sum_{s^{t+1} | s^t} \frac{\partial}{\partial a_i} [u (c_{i,t+1} (s^{t+1})) - v (a_{i,t+1} (s^{t+1}))] \pi (s_{t+1} | s_{t}, a_{t} (s^t)) \\
/ \lambda_{it} (s^t) : & \quad - v (a_{i,t+j} (s^{t+j})) \Pi (s^{t+j} | s^t, a^{t+j-1} (s^{t+j-1})) = 0 \\
\end{align*}
\]

We can restate all the above equations in only one endogenous state variable \( \theta_t (s^t) \equiv \frac{\phi_{2it}(s^t)}{\phi_{it}(s^t)} \), by using the homogeneity properties of policy functions and value function. For a generic variable \( x_{it} (s^t) \), define \( \tilde{x}_{it} (s^t) = \frac{x_{it}(s^t)}{\phi_{it}(s^t)} \), thus:

\[
\begin{align*}
/ c_{it} (s^t) : & \quad \frac{u' (c_{it} (s^t))}{u' (c_{2it} (s^t))} = \theta_t (s^t) \\
/ a_{it} (s^t) : & \quad u' (a_{it} (s^t)) + \lambda_{it} (s^t) u'' (a_{it} (s^t)) = \beta \sum_{s^{t+1} | s^t} \left( 1 + \tilde{\lambda}_{it} (s^t) \frac{\pi_{ai} (s_{i,t+1} | s_{it}, a_{it} (s^t))}{\pi (s_{i,t+1} | s_{it}, a_{it} (s^t))} \right) \times \\
&\times \pi_{ai} (s_{i,t+1} | s_{it}, a_{it} (s^t)) \pi (s_{2t+1} | s_{2t}, a_{2t} (s^t)) \tilde{J} (s_t, \theta_t (s^t)) + \\
&+ \beta \tilde{\lambda}_{it} (s^t) \sum_{s^{t+1} | s^t} \frac{\partial}{\partial a_i} [u (c_{i,t+1} (s^{t+1})) - v (a_{1,t+1} (s^{t+1}))] \pi (s_{t+1} | s_{t}, a_{t} (s^t)) \\
\end{align*}
\]
\[ /a_{2t}(s^t) : \quad \theta_t(s^t) v'(a_{2t}(s^t)) + \theta_t(s^t) \tilde{\lambda}_{2t}(s^t) v''(a_{2t}(s^t)) = \beta \sum_{s^{t+1}\mid a^t} \left( 1 + \tilde{\lambda}_{1t}(s^t) \frac{\pi_{a_i}(s_{1, t+1} \mid s_{1t}, a_{1t}(s^t))}{\pi(s_{1, t+1} \mid s_{1t}, a_{1t}(s^t))} \right) \times \]

\[ \times \pi(s_{1, t+1} \mid s_{1t}, a_{1t}(s^t)) \pi_{a_2}(s_{2, t+1} \mid s_{2t}, a_{2t}(s^t)) \tilde{J}(s_t, \theta_t(s^t)) + \]

\[ + \beta \theta_t(s^t) \tilde{\lambda}_{2t}(s^t) \sum_{s^{t+1}\mid a^t} \frac{\partial}{\partial a_2} \left( \frac{\pi_{a_2}(s_{2, t+1} \mid s_{2t}, a_{2t}(s^t))}{\pi(s_{2, t+1} \mid s_{2t}, a_{2t}(s^t))} \right) \left[ u(c_{2, t+1}(s^{t+1})) - u(a_{2, t+1}(s^{t+1})) \right] \pi(s_{t+1} \mid s_t, a_t(s^t)) \]

\[ /\lambda_{it}(s^t) : \quad -v'(a_{it}(s^t)) + \sum_{j=1}^{\infty} \sum_{s^t+j\mid a^t} \beta_j \pi_{a_i}(s_{i, t+1} \mid s_{it}, a_{it}(s^t)) \pi(s_{i, t+1} \mid s_{it}, a_{it}(s^t)) \left[ u(c_{i, t+j}(s^{t+j})) - u(a_{i, t+j}(s^{t+j})) \right] \]

\[ \pi(s_{i, t+j} \mid s^t, a^{t+j-1}(s^{t+j-1})) = 0 \]
Appendix D: Bond holdings

6.6 Repeated moral hazard

We can define bond holdings recursively as:

$$b_t(s^t) = -E_t^a \sum_{j=1}^{\infty} \beta^j (y_{t+j} - c_{t+j}) =$$

$$= -E_t^a \sum_{j=1}^{\infty} \beta^j \{ (y_{t+j} - c_{t+j}) + \phi_{t+j} [u(c_{t+j}) - v(a_{t+j})] - \lambda_{t+j} v'(a_{t+j}) \} +$$

$$+ E_t^a \sum_{j=1}^{\infty} \beta^j \{ \phi_{t+j} [u(c_{t+j}) - v(a_{t+j})] - \lambda_{t+j} v'(a_{t+j}) \}$$

$$= -\beta E_t^a J(y_{t+1}, \phi_{t+1}) + E_t^a \sum_{j=1}^{\infty} \beta^j \{ \phi_{t+j} [u(c_{t+j}) - v(a_{t+j})] - \lambda_{t+j} v'(a_{t+j}) \}$$

$$= -\beta E_t^a J(y_{t+1}, \phi_{t+1}) + E_t^a \sum_{j=1}^{\infty} \beta^j \{ \phi_{t+j} [u(c_{t+j}) - v(a_{t+j})] \}$$

$$- E_t^a \sum_{j=1}^{\infty} \beta \lambda_{t+j} E_{t+1}^a \sum_{k=0}^{\infty} \beta^k \frac{\pi a}{\pi (a_{t+j+1})} [u(c_{t+j+k+1}) - v(a_{t+j+k+1})]$$

$$= -\beta E_t^a J(y_{t+1}, \phi_{t+1}) + E_t^a \sum_{j=1}^{\infty} \beta^j \{ \phi_{t+j} [u(c_{t+j}) - v(a_{t+j})] \}$$

$$- E_t^a \sum_{j=1}^{\infty} \beta \lambda_{t+j} \frac{\pi a}{\pi (a_{t+j})} U(s_{t+j+1}, \phi_{t+j+1})$$

and notice that

$$\phi^*_t U(s_t, \phi^*_t) = \phi^*_t [u(c^*_t) - v(a^*_t)] + \phi^*_t \beta E_t^a U(s_{t+1}, \phi^*_t)$$

$$= \phi^*_t [u(c^*_t) - v(a^*_t)] + \beta E_t^a \phi^*_t \frac{\phi^*_t+1}{\phi^*_t} U(s_{t+1}, \phi^*_t+1)$$

$$= \phi^*_t [u(c^*_t) - v(a^*_t)] + \beta E_t^a \phi^*_t \frac{\phi^*_t+1}{\phi^*_t+1} U(s_{t+1}, \phi^*_t+1)$$

$$= \phi^*_t [u(c^*_t) - v(a^*_t)] + \beta E_t^a \phi^*_t+1 U(s_{t+1}, \phi^*_t+1) =$$

$$= \phi^*_t [u(c^*_t) - v(a^*_t)] + E_t^a \sum_{j=1}^{\infty} \beta \phi^*_t [u(c^*_{t+j}) - v(a^*_{t+j})]$$
due to homogeneity of degree zero of the policy functions and of \(U(s, \cdot)\). Therefore

\[
\begin{align*}
\Delta t (s^t) &= -\beta E_t^a \mathcal{J} (y_{t+1}, \phi_{t+1}) + \beta E_t^a \phi_{t+1} U (y_{t+1}, \phi_{t+1}) \\
&= -E_t^a \sum_{j=1}^{\infty} \beta^j \lambda_{t+j} \frac{\pi_a (a_{t+j})}{\pi(a_{t+1})} U (y_{t+j+1}, \phi_{t+j+1})
\end{align*}
\]

by Abel’s formula

\[
\begin{align*}
&= -\beta E_t^a \mathcal{J} (y_{t+1}, \phi_{t+1}) + \beta E_t^a \phi_{t+1} U (y_{t+1}, \phi_{t+1}) \\
&= -E_t^a \sum_{j=1}^{\infty} \beta^j (\phi_{t+j+1} - \phi_{t+1}) \left[ u (c_{t+j+1}^*) - v (a_{t+j+1}^*) \right] \\
&= -\beta E_t^a \mathcal{J} (y_{t+1}, \phi_{t+1}) + \beta E_t^a \phi_{t+1} U (y_{t+1}, \phi_{t+1}) \\
&= -E_t^a \sum_{j=1}^{\infty} \beta^j (\phi_{t+j+1} - \phi_{t+1}) \left[ u (c_{t+j+1}^*) - v (a_{t+j+1}^*) \right] \\
&= -\beta E_t^a \mathcal{J} (y_{t+1}, \phi_{t+1}) + \beta E_t^a \phi_{t+1} U (y_{t+1}, \phi_{t+1})
\end{align*}
\]

which can be rewritten as

\[
\begin{align*}
\Delta t (s^t) &= -\beta E_t^a \mathcal{J} (y_{t+1}, \phi_{t+1}) + \beta E_t^a \phi_{t+1} U (y_{t+1}, \phi_{t+1}) \\
&= -E_t^a \lambda_{t+1} \frac{\pi_a (a_{t+1})}{\pi(a_{t+1})} \left[ u (c_{t+1}^*) - v (a_{t+1}^*) \right]
\end{align*}
\]

where the second line is due to the optimality of the contract.

### 6.7 Hidden assets

Starting from the previous result, in this case we can write

\[
\begin{align*}
\Delta t (s^t) &= -E_t^a \sum_{j=1}^{\infty} \beta^j (y_{t+j} - c_{t+j}) = \\
&= -\beta E_t^a \mathcal{J} (y_{t+1}, \phi_{t+1}) + \beta E_t^a \phi_{t+1} U (y_{t+1}, \phi_{t+1}) - E_t^a \lambda_{t+1} \frac{\pi_a (a_{t+1})}{\pi(a_{t+1})} \left[ u (c_{t+1}^*) - v (a_{t+1}^*) \right] \\
&= -\beta E_t^a \mathcal{J} (y_{t+1}, \phi_{t+1}) + \beta E_t^a \phi_{t+1} U (y_{t+1}, \phi_{t+1}) - E_t^a \lambda_{t+1} \frac{\pi_a (a_{t+1})}{\pi(a_{t+1})} \left[ u (c_{t+1}^*) - v (a_{t+1}^*) \right] \\
&= -\beta E_t^a \mathcal{J} (y_{t+1}, \phi_{t+1}) + \beta E_t^a \phi_{t+1} U (y_{t+1}, \phi_{t+1}) - E_t^a \lambda_{t+1} \frac{\pi_a (a_{t+1})}{\pi(a_{t+1})} \left[ u (c_{t+1}^*) - v (a_{t+1}^*) \right] = 0 \text{ by definition}
\end{align*}
\]

\[
\begin{align*}
&= -\beta E_t^a \mathcal{J} (y_{t+1}, \phi_{t+1}) + \beta E_t^a \phi_{t+1} U (y_{t+1}, \phi_{t+1}) - E_t^a \lambda_{t+1} \frac{\pi_a (a_{t+1})}{\pi(a_{t+1})} \left[ u (c_{t+1}^*) - v (a_{t+1}^*) \right] \\
&= +E_t^a \zeta_{t+1} u' (c_{t+1})
\end{align*}
\]
Figures

Figure 1: Pure moral hazard, policy functions

Figure 2: Pure moral hazard, policy functions, cont.
Figure 3: Pure moral hazard, average over 50000 independent simulations

Figure 4: Pure moral hazard, average over 50000 independent simulations, cont.
Figure 5: Pure moral hazard, Pareto frontier
Figure 6: Moral hazard with hidden assets, policy functions

Figure 7: Moral hazard with hidden assets, policy functions, cont.
Figure 8: Moral hazard with hidden assets, average over 50000 independent simulations

Figure 9: Moral hazard with hidden assets, average over 50000 independent simulations, cont.
Figure 10: Moral hazard with hidden assets, Pareto frontier
Figure 11: Risk sharing with moral hazard, policy functions (2 agents)

Figure 12: Risk sharing with moral hazard, policy functions (2 agents), cont.
Figure 13: Risk sharing with moral hazard, sample path (2 agents)

Figure 14: Risk sharing with moral hazard, sample path (2 agents), cont.
Figure 15: Risk sharing with moral hazard, Pareto frontier (2 agents)
Figure 16: Risk sharing with moral hazard and capital accumulation, average over 50000 independent simulations

Figure 17: Risk sharing with moral hazard and capital accumulation, average over 50000 independent simulations (cont)
Figure 18: Risk sharing with moral hazard and capital accumulation, Pareto frontier