INNOCENZO GASPARINI INSTITUTE FOR ECONOMIC RESEARCH

Institutional Members: CEPR, NBER and Università Bocconi

WORKING PAPER SERIES

Contractive Dual Methods for Incentive Problems

Matthias Messner, Nicola Pavoni, Christopher Sleet

Working Paper n. 466

This Version: July, 2012

IGIER – Università Bocconi, Via Guglielmo Röntgen 1, 20136 Milano –Italy
http://www.igier.unibocconi.it

Contractive Dual Methods for Incentive Problems

Matthias Messner* Nicola Pavoni† Christopher Sleet‡

This version: July 2012

Abstract

Several recent papers have proposed recursive Lagrangian-based methods for solving dynamic contracting problems. These methods give rise to Bellman operators that incorporate either a dual inf-sup or a saddle point operation. We give conditions that ensure the Bellman operator implied by a dual recursive formulation is contractive.

JEL codes: C61, C73, D82, E61.

Keywords: Dynamic Contracts, Duality, Dynamic Programming, Contraction Mapping Theorem.

1 Introduction

Dynamic incentive models have received widespread application in recent research. Central to these models are constraints that ensure uncommitted or privately informed players are motivated to take prescribed courses of actions. The forward-looking nature of these incentive constraints disrupts the principle of optimality and complicates the task of finding an optimal plan or contract. Two approaches have been proposed for resolving this difficulty and obtaining a recursive formulation. One augments the problem with additional promise-keeping constraints; the other perturbs the objective with dual state variables introduced via the Lagrangian. Both approaches lead to Bellman operators that are, potentially, the bases of constructive value iteration procedures. However, the development of these procedures presents challenges. This paper is concerned with the second

*Department of Economics and IGIER, Bocconi University, I-20136, Milan, Italy; matthias.messner@unibocconi.it.

†Department of Economics and IGIER, Bocconi University, I-20136, Milan, Italy; Department of Economics, EUI, I-50133 Firenze, Italy; CEPR and IFS, London; pavoni.nicola@gmail.com.

‡Corresponding author. Tepper School of Business, Carnegie Mellon University, Pittsburgh PA 15217; csleet@andrew.cmu.edu. Phone: +44 412 268 9938. Fax +44 412 268 2268.

(dual) approach. It gives conditions that ensure the associated Bellman operator is a contraction on a suitable space of functions. The contraction mapping theorem then ensures convergent value iteration and provides rates of convergence.

Messner, Pavoni, and Sleet (2011a) and (2011b) develop dual recursive methods for dynamic incentive problems. These methods lead to a dynamic programming formulation in which Lagrange multipliers from current incentive constraints act as controls and accumulated multipliers from past periods act as state variables.[1] The associated dual Bellman operator updates candidate value functions via a family of "inf-sup" operations over Lagrange multipliers and choice variables. Specifically, at each dual state and current constraint Lagrange multiplier, "inner" supremum operations are performed over current choices. The resulting current (indirect) objective is combined with a value function over future dual states and a law of motion for such states to give an objective over current dual states and multipliers. At each current dual state, an outer infimum operation over multipliers gives the updated value function.

Under appropriate concavity conditions, it is relatively straightforward to show that the optimal (dual) value function satisfies a dual Bellman equation, i.e. is a fixed point of a dual Bellman operator. However, relative to the more standard problems described in Stokey, Lucas, and Prescott (1989), there are several difficulties in recovering the optimal value function from the dual Bellman operator. First, the optimal dual value function may be infinite-valued at some dual states. As an initial step, it is necessary to identify the (endogenous) state space on which the value function is finite. In this paper, we focus on problems in which the optimal dual value function is everywhere real-valued. This covers many incentive problems with bounded agent payoffs. Further discussion of problems with a more restricted, but known state space is relegated to an appendix. Second, even on this state space, the value function is unbounded (in the usual sup-norm sense). To apply the contraction mapping approach, it is necessary to identify a complete metric space of functions containing the optimal value function and on which the dual Bellman is a contractive self-map. Following Wessels (1977), a common procedure is to obtain a set of functions $\mathscr{F}$ containing the optimal value function, closed and bounded with respect to a weighted sup norm[2] and on which the Bellman is contractive. However, the use of weighted sup-norms requires that the continuation state variables and, hence, the continuation value function cannot grow "too much" on the graph of the constraint corre-

[1] In many applications, the dual state variables are shadow prices on primal promise-keeping constraints and can be interpreted as agent Pareto weights.

[2] A weighted sup norm on a set of functions $\mathscr{F}$ with common domain X is a function $\|\cdot\|_w : \mathscr{F} \to \mathbb{R}$ of the form $\|f\|_w = \sup_X \left|\frac{f(x)}{w(x)}\right|$ for some $w : X \to \mathbb{R}_{++}$.

spondence. Since the basic dual Bellman operator permits the choice of any non-negative multiplier, this requirement is not usually met. Cole and Kubler (2010) make progress by obtaining additional non-binding constraints on multipliers and, hence, continuation states, that enable the weighted norm approach to be applied. The restrictions on primitives for these additional constraints to be non-binding appear strong. We proceed quite differently. Following an idea of Rincon-Zapatero and Rodriguez-Palermo (2003), we exploit the concavity of the Bellman operator to establish contractivity.[3] Key to the approach is the derivation of bounding value functions with desired properties. We show through examples that such bounding functions are often available. A final difficulty is that the unboundedness and, hence, non-compactness of the set of feasible multipliers disrupts the application of the Theorem of the Maximum. However, it is easy to show that the optimal value function is concave. When it is everywhere real-valued as well, appeals can be made to the continuity properties of concave functions to establish continuity of the optimal value function. When the optimal value function is not everywhere real-valued, then convexity does not ensure continuity. In an appendix, we show how level-boundedness of the objective and constraints can be used to establish lower semicontinuity of value functions in this case.

Under standard conditions for the equivalence of primal and dual values, dual dynamic programming problems provide necessary, but not sufficient conditions for optimal plans (i.e. for solutions to an original incentive problem). Despite this, in many applications the law of motion for cumulative multipliers is constrained to be a martingale, sub-martingale or iterated function system. Long run properties of these processes can be derived and, hence, long run properties of optimal plans inferred.

The paper proceeds as follows. After a brief review of related literature, Section 3 outlines the family of incentive constrained problems we consider. Section 4 applies the dual recursive method to this family and derives the associated Bellman equation. Sufficient conditions for the associated Bellman operator to be contractive are given in Section 5. Applications to private information and limited commitment problems are provided in Examples 1 and 2. Implications for policies are discussed in Section 7.

1.1 Related literature

The idea of using a recursive decomposition of the Lagrangian to solve an optimal control problem originates with Marcet and Marimon (1999) (revised: Marcet and Marimon (2011)). Their approach is distinct from ours in that they decompose a saddle point

[3] As opposed to verifying discounting.

rather than a dual problem. Messner, Pavoni, and Sleet (2011a) develops a recursive dual approach in both time additive and non-additive settings; Messner, Pavoni, and Sleet (2011b) relates the promise-keeping and recursive dual approaches in infinite horizon concave settings. It exploits the monotonicity of these operators to obtain iterative procedures for computing value functions. Cole and Kubler (2010) provide a clever extension of dual recursive-type methods that allows them to recover optimal policies in wide range of concave settings, including ones concavified by the use of lotteries.

There has been much focus in the dynamic programming literature on developing techniques applicable to sup-unbounded problems. Approaches based on the use of sup weight-norms were developed by Wessels (1977) and applied to economic problems by Boyd (1990) and Duran (2000).[4] Rincon-Zapatero and Rodriguez-Palermo (2003) proposed a localized approach to building a complete metric space and verifying the contractive properties of the Bellman operator.[5] These authors also show how the concavity property of the Bellman operator can be used to establish contractivity.

Several previous attempts have been made to establish the contractivity of related Bellman operators. Marcet and Marimon (1999) and Mele (2010) pursue a version of the weighted contraction mapping approach for homogenous problems described in Stokey, Lucas, and Prescott (1989). This involves verifying that the Bellman operator satisfies Blackwell's conditions on an enlarged space of functions bounded with respect to a weighted sup norm. Unfortunately, as mentioned, the combination of an unbounded candidate value function and an unbounded constraint set make this difficult and the proofs provided by, say, Mele (2010) contain errors. Cole and Kubler (2010) does contain a successful proof in this direction, but under quite restrictive conditions[6].

2 A Simple Example

To make the subsequent analysis concrete, we briefly outline a simple one sided limited commitment example. A committed principal shares risk with an uncommitted agent. Both live for an infinite number of periods $t \in \mathbb{N}$. Let $\{s_t\}_{t=1}^{\infty}$ be a shock process taking values in a finite set $\mathbb{S} = \{1, \ldots, S\}$ and evolving according to a Markov transition Q from $s_0 \in \mathbb{S}$. Let the endowment of goods in each current state be given by $\omega : \mathbb{S} \to \mathbb{R}_{++}$. A consumption process for the agent $\{c_t\}_{t=1}^{\infty}$, $c_t : \mathbb{S}^t \to \mathbb{R}_+$, is resource-

[4] For a textbook treatment, see Hernandez-Lerma and Laserre (1999).

[5] This paper has been influential, with variations (and a correction) contained in Matkowski and Nowak (2011), Rincon-Zapatero and Rodriguez-Palermo (2009) and da Rocha and Vailakis (2010).

[6] In fairness, establishing such a result is not the focus of Cole and Kubler (2010)

feasible if for all $t \in \mathbb{N}$, $s^t \in \mathbb{S}^t$, $c_t(s^t) \in [0, \omega(s_t)]$. Given c^∞, the principal receives the residual process, $\forall t, s^t \in \mathbb{S}^t$, $\omega(s_t) - c_t(s^t)$. Conditional on the first period shock, the agent and the principal value resource-feasible consumption processes c^∞ according to: $U^c(s', c^\infty) = \sum_{t=1}^{\infty} \beta^{t-1} E[v(c_t(s^t)) | s_1 = s']$ and $\sum_{t=1}^{\infty} \beta^{t-1} E[v(\omega_t(s_t) - c_t(s^t)) | s_1 = s']$, respectively, where $\beta \in (0,1)$ and $v : \mathbb{R}_+ \to \mathbb{R}$ is increasing. We impose the following assumption.

Assumption 1. $v : \mathbb{R}_+ \to \mathbb{R}$ *is continuous, increasing and strictly concave.*

c^∞ is incentive-feasible if it is resource feasible and it provides the agent with utility in excess of her outside option B:

$$\forall t, s^t, \quad v(c_t(s^t)) + \beta \sum_{s' \in \mathbb{S}} U^c(s', c^\infty_{t+1}(s^t)) Q(s_t, s') \geq B(s_t),$$

where $c^\infty_{t+1}(s^t) = \{c_{t+j}(s^t, \cdot)\}_{j=1}^{\infty}$ and $B : \mathbb{S} \to \mathbb{R}$ satisfies: $B(s) = b(s) + \beta \sum_{s' \in \mathbb{S}} B(s') Q(s, s')$, $b : \mathbb{S} \to \mathbb{R}$.

Assumption 2. $\forall s$, $b(s) \in [v(0), v(\omega(s))]$.

It is convenient to re-express the problem in terms of agent utility net of the outside option. Given c^∞ the associated net utility plan $y^\infty = \{y_t\}_{t=1}^{\infty}$ is defined as for all t, s^t, $y_t(s^t) = v(c_t(s^t)) - b(s_t)$. Let $\kappa = v^{-1}$ and $C(s) = \{y : \omega(s) - \kappa(y + b(s)) \geq 0\}$. By Assumptions 1 and 2, C is non-empty and compact-valued. $\Omega_0 = \{y^\infty | \forall t, y_t(\cdot, s_t) \in C(s_t)\}$ gives the set of resource-feasible net utility plans. A plan $y^\infty \in \Omega_0$ is valued according to:

$$U(s', y^\infty) = \sum_{t=1}^{\infty} \beta^{t-1} E[y_t(s^t) | s_1 = s']$$

by the agent (conditional on period 1 shock s' and net of the outside option) and according to:

$$F(s, y^\infty) = \sum_{t=1}^{\infty} \beta^{t-1} E[f(s_t, y_t(s^t)) | s_0 = s], \quad f(s, y) = v(\omega(s) - \kappa(y + b(s))),$$

by the principal (conditional on the period 0 shock s). A net utility plan in Ω_0 is incentive-feasible if it delivers a non-negative net utility to the agent for all t, s^{t-1} and s,

$$y_t(s^{t-1}, s) + \beta \sum_{s' \in \mathbb{S}} U(s', y^\infty_{t+1}(s^{t-1}, s)) Q(s, s') \geq 0. \tag{1}$$

Now, $M = S$ is the number of per period constraints. Let Ω_1 denote the set of incentive-feasible (net utility) plans. The principal's problem is:

$$\sup_{\Omega_1} F(s, y^\infty). \tag{2}$$

We associate with this a family of perturbed problems, $\forall \mu \in \mathbb{R}$,

$$V^*(s,\mu) = \sup_{\Omega_1} F(s, y^\infty) + \mu \sum_{s' \in \mathbb{S}} U(s', y^\infty) Q(s,s'). \tag{3}$$

where $V^* : \mathbb{R} \to \mathbb{R} \cup \{\infty\}$ and the principal's original problem (2) is recovered by setting $\mu = 0$. Next absorb the first period incentive constraints into a Lagrangian:

$$\begin{aligned}\mathscr{L}_s(y^\infty, \eta; \mu) = F(s, y^\infty) + \mu \sum_{\mathbb{S}} U(s', y^\infty) Q(s, s') \\ + \sum_{s'} \eta^{s'} Q(s,s') \Big\{ y_1(s') + \delta \sum_{s'' \in \mathbb{S}} U(s'', y_2^\infty(s')) Q(s', s'') \Big\},\end{aligned}$$

where $\eta^{s'} Q(s,s')$ is the multiplier on the s'-th first period constraint. The optimization in (3) may be rewritten in primal sup-inf form as:

$$V^*(s,\mu) = \sup_{\Omega_2} \inf_{\mathbb{R}^M_+} \mathscr{L}_s(y^\infty, \eta; \mu), \tag{4}$$

where constraints for periods after the first remain within the constraint set Ω_2. A dual problem is obtained by interchanging the supremum and infimum operations in (4). In general the optimal value from the dual need not equal that from the primal. However, in this case, under Assumptions 1 and 2, these values are equal and:

$$V^*(s,\mu) = \inf_{\mathbb{R}^M_+} \sup_{\Omega_2} \mathscr{L}(y^\infty, \eta; \mu). \tag{5}$$

Grouping terms involving $y_1(s)$ together gives:

$$\begin{aligned}V^*(s,\mu) = \inf_{\mathbb{R}^M_+} \sup_{\Omega_2} \sum_{\mathbb{S}} \Big\{ f(s', y_1(s')) + \left(\mu + \eta^{s'}\right) y_1(s) \\ + \beta \left(F(s'', y_2^\infty(s')) + \left(\mu + \eta^{s'}\right) U(s'', y_2^\infty(s')) \right) \Big\} Q(s,s').\end{aligned}$$

Thus, using the definition of f and V^*:

$$V^*(s,\mu) = \inf_{\mathbb{R}^M_+} \sum_{\mathbb{S}} \left[\nu_{s'}(\mu,\eta) + \beta V^*\left(s', \mu'_{s'}(\mu,\eta)\right)\right] Q(s,s'),$$

with $\nu_{s'}(\mu,\eta) = \sup_{y \in C(s')}(\mu + \eta^{s'})y + f(s',y)$ and weight updating formula $\mu'_{s'}(\mu,\eta) = \mu + \eta^{s'}$. It follows that V^* satisfies the Bellman equation $V^* = \mathscr{D}(V^*)$, where for $V : \mathbb{S} \times \mathbb{R} \to \mathbb{R} \cup \{\infty\}$ and $(s,\mu) \in \mathbb{S} \times \mathbb{R}$:

$$\mathscr{D}(V)(s,\mu) = \inf_{\mathbb{R}^M_+} \sum_{\mathbb{S}} \left[\nu_{s'}(\mu,\eta) + \beta V\left(s', \mu'_{s'}(\mu,\eta)\right)\right] Q(s,s').$$

We seek to show the contractivity of $\mathscr{D}$ on a suitable space of functions for this and other related problems.

3 Incentive-constrained problems

We outline a framework that accommodates many incentive-constrained problems. Let $\mathbb{I} = \{1,\ldots,I\}$ denote a finite set of agents and $\mathbb{S} = \{1,\ldots,S\}$ a finite set of shocks. A shock s may include components that affect all agents and/or components that are idiosyncratic to specific agents. In the latter case, these components may be common knowledge or privately observed by the affected agents. Identify time with the natural numbers $t \in \mathbb{N}$. Let s_t be an $\mathbb{S}$-valued random variable describing the shock in period t and $s^t = (s_1,\ldots,s_t)$ an $\mathbb{S}^t$-valued random variable describing shock histories up to t. Let Q denote a Markov transition for shocks, which, given our finiteness assumption, we identify with a $S \times S$-matrix of transition probabilities $\{Q(s,s')\}$, and $E[\cdot|s]$ the conditional expectation operator induced by Q and a current shock s.

Let $Y(i)$, $i \in \mathbb{I}$, denote a set of actions for the i-th agent, $Y = \prod_{i \in \mathbb{I}} Y(i)$ a set for the population and $C : \mathbb{S} \to 2^Y \backslash \emptyset$ a non-empty valued correspondence mapping shocks to action sets. An action plan is a sequence $y^\infty = \{y_t\}$ with $y_t : \mathbb{S}^t \to Y$ and $y_t(s^t) \in C(s_t)$. Let u^i : Graph $C \to \mathbb{R}$, $i \in \mathbb{I}$, give the i-th agent's per period payoff as a function of the current shock and action and let $\delta \in [0,1)$ denote the agent's discount factor. Given an initial shock s and action plan $y^\infty = \{y_t\}$, the i-th agent's lifetime payoff is:

$$U^i(s,y^\infty) := \liminf_{T\to\infty} \sum_{t=1}^{T} \delta^{t-1} E[u^i(s_t, y_t(s^t))|s_1 = s].$$

Let:

$$\Omega_0 = \left\{ y^\infty \,\middle|\, \begin{array}{l} \forall t, y_t(s^t) \in C(s_t), \\ \forall i \in \mathbb{I}, s \in \mathbb{S}, U^i(s, y^\infty) = \lim_{T\to\infty} \sum_{t=1}^{T} \delta^{t-1} E[u^i(s_t, y_t(z^t))|s] \in \mathbb{R} \end{array} \right\}.$$

We restrict attention to plans in Ω_0. Let $y_t^\infty(s^{t-1})$ denote the continuation of plan y^∞ following the shock history s^{t-1}.

The restrictions in Ω_0 are supplemented with forward-looking incentive constraints. These constraints guarantee that a given plan is optimal for an agent relative to some set of feasible deviations. A key piece of constraint structure common to many examples, including the one sided commitment example of Section 2 and others considered later, is linearity of incentive constraints in continuation payoffs. In other dimensions incentive constraints are problem specific: in limited commitment problems they are state specific, while in hidden information problems they run across states. This leads us to write down a constraint structure that is additive across time and states and linear in agent continuation payoffs, but is otherwise general. The functions defining incentive constraints are constructed from three elements: a set of labels $\mathbb{M}$, a payoff function u^m and a family of continuation payoff weights $\{q^{m,i}_{s,s'}\}$. The labels $m \in \mathbb{M} := \{1, \ldots, M\}$ identify the constraints and encode the identity of a prescribed action and a deviation from this action. The current payoff function $u^m : \text{Graph}\, C \to \mathbb{R}$ describes how the action(s) taken by agents in each current shock state enter the m-th constraint. Agent continuation payoffs enter the incentive constraints linearly. The terms $q = \{q^{m,i}_{s,s'}\}$ weight them: $q^{m,i}_{s,s'}$ describes how the continuation payoff of the i-th agent following the s-th action today and the s'-th shock tomorrow is weighted in the m-th constraint. Collecting elements together, the following definition is obtained.

Definition 1. A plan $y^\infty = \{y_t\}_{t=1}^\infty \in \Omega_0$ is *incentive-compatible* if for each $t \in \mathbb{N}$, $s^{t-1} \in \mathbb{S}^{t-1}$, $G(y_t^\infty(s^{t-1})) \geq 0$, where $G : \Omega_0 \to \mathbb{R}^M$ and

$$G(y^\infty) = \{G_m(y^\infty)\}_{m\in\mathbb{M}} = \left\{ \sum_{\mathbb{S}} \left[u^m(s, y_1(s)) + \delta \sum_{\mathbb{I}} \sum_{\mathbb{S}} U^i(s', y_2^\infty(s)) q^{m,i}_{s,s'} \right] \right\}_{m\in\mathbb{M}} \geq 0. \tag{6}$$

Let Ω_1 denote the set of incentive-compatible plans. Let Ω_2 denote the set of plans whose continuations are incentive compatible, i.e. y^∞ such that each $y_2^\infty(s) \in \Omega_1$.

It is easy to see that the one sided limited commitment example in Section 2 has (and indeed many other examples have) incentive constraints satisfying (6). For example, the constraints in the limited commitment problem (1) may be obtained from (6) by setting, for each current shock $m \in \mathbb{S}$,

$$u^m(s,y) = \begin{cases} y & \text{if } s = m \\ 0 & \text{if } s \neq m \end{cases} \quad \text{and} \quad q^m_{s,s'} = \begin{cases} Q(s,s') & \text{if } s = m \\ 0 & \text{if } s \neq m. \end{cases} \tag{7}$$

Thus, actions and payoffs not associated with the m-th shock are zeroed out of the m-th constraint.

An important special case of (6) occurs when the weights $q^{m,i}_{s,s'}$ can be decomposed as:

$$q^{m,i}_{s,s'} = q^{m,i}_s Q(s,s') \;\; (s,s') \in \mathbb{S}^2, \tag{8}$$

for some $\{q^{m,i}_s\}$. In this case, the incentive function G can be re-expressed as:

$$G(y^\infty) = \left\{ \sum_{\mathbb{S}} \left[u^m(s, y_1(s)) + \delta \sum_{\mathbb{I}} q^{m,i}_s \sum_{\mathbb{S}} U^i(s', y^\infty_2(s)) Q(s,s') \right] \right\}_{m \in \mathbb{M}}. \tag{9}$$

with the $q^{m,i}_s$ terms factored through the summation over s' values. Thus, the continuation plan $y^\infty_2(s)$ affects the constraint functions only insofar as it affects the conditional expected payoff $\sum_{\mathbb{S}} U^i(s', y^\infty_2(s))Q(s,s')$. This constraint structure affords a considerable simplification of the analysis and is quite common in applications. For example, in the one sided commitment problem, the weights $q^m_{s,s'}$ may be factored as $q^m_s Q(s,s')$, where $q^m_s = 1$ if $s = m$ and 0 otherwise.

Let f : Graph $C \to \mathbb{R}$ denote the principal's per period payoff as a function of the current shock and action. Together f, a discount factor $\beta \in [\delta, 1)$ and Q give the principal's lifetime payoff, $F : \mathbb{S} \times \Omega_1 \to \overline{\mathbb{R}}$, where:

$$F(s, y^\infty) := \liminf_{T \to \infty} \sum_{t=1}^{T} \beta^{t-1} E\left[f(s_t, y_t) \middle| s_0 = s \right].$$

Note that this objective is conditional on the date 0 not the date 1 shock. The basic dynamic incentive problem is then:

$$\sup_{\Omega_1} F(s, y^\infty). \tag{10}$$

Assumptions 3 and 4 collect basic conditions imposed on the primitives of this problem. Here and throughout the paper if a function $g : \mathbb{S} \times \mathbb{R}^n \to \mathbb{R}$ is such that each $g(s, \cdot)$ satisfies a property (e.g. convexity or continuity), then we will simply say that g possesses the property.

Assumption 3 (Constraints)**.** *The following conditions are imposed on constraints.*

(i) $C : \mathbb{S} \to 2^Y \backslash \emptyset$ *is convex-valued.*

(ii) Each u^i *: Graph* $C \to \mathbb{R}$ *and* u^m *: Graph* $C \to \mathbb{R}$ *is concave and continuous.*

(iii) For each m *and* s, $\sum_{\mathbb{I}} \sum_{\mathbb{S}} U^i(s', \cdot) q^{m,i}_{s,s'}$ *is concave.*

(iv) Ω_1 *is non-empty.*

(v) (Slater) $\exists y^\infty \in \Omega_2$ *with* $G(y^\infty) > 0$.

Assumption 3 (ii) and (iii) ensure that G is a concave function. Assumption 3 (iii) holds if, in addition to Assumption 3 (ii), each $q^{m,i}_{s,s'}$ is non-negative (as is typically the case in limited commitment problems) or each u^i is affine (as is the case in hidden information problems when agent's have separable preferences and utilities are made the choice variable). These conditions coupled with the non-emptiness condition (iv) and the Slater condition (v) ensure that standard duality results can be invoked. They are easy to check, but stronger than needed. Assumption 4 restricts the objective function in (10).

Assumption 4 (Objective)**.** *The following conditions are imposed on the objective.*

(i) Each $f : C \to \mathbb{R}$ *is concave, continuous and bounded above.*

(ii) f, β, Q *and* Ω_1 *are such that for all* $s \in \mathbb{S}$ *and* $y^\infty \in \Omega_1$,

$$F(s, y^\infty) = \lim_{T \to \infty} \sum_{t=1}^{T} \beta^{t-1} E\left[f(s_t, y_t) \middle| s_0 = s \right] \in \mathbb{R}.$$

Condition Assumption 4 may sometimes require that Ω_0 (and, hence, Ω_1) is further restricted to ensure the principal's objective is real-valued (and not $-\infty$-valued).

4 The Dual Recursive method

Actions associated with a given period appear in the incentive constraints of preceding periods. Consequently, the family of choice problems (10) is time inconsistent and the principle of optimality is disrupted.[7] The Dual Recursive method resolves this difficulty by augmenting the *objectives* from the problems in (10) with additional terms that capture the impact of past incentive provision. Lagrangians that incorporate current incentive

[7] A continuation plan provides incentives in past periods. When it is time for the plan to be implemented, these incentive benefits are sunk.

constraints are constructed for these augmented problems. Multipliers from these constraints are used to update the continuation objective and, hence, enforce consistency of this objective with past incentive provision. Time consistency is restored.

4.1 Generalized problem

Agents are partitioned into two groups: $\mathbb{I}_1$ and $\mathbb{I}_2$. $\mathbb{I}_1$ consists of those agents for whom each $q^{m,i}$ satisfies (8); $\mathbb{I}_2$ contains the remaining agents. The *perturbed incentive problem* is:

$$V^*(s,\mu) = \sup_{\Omega_1} F(s,y^\infty) + \sum_{\mathbb{I}_1} \mu^i \sum_{\mathbb{S}} U^i(s',y^\infty)Q(s,s') + \sum_{\mathbb{I}_2}\sum_{\mathbb{S}} \mu^i_{s'} U^i(s',y^\infty)Q(s,s'). \quad (11)$$

We refer to the μ^i terms as *weights* and collect them into $\mu := \{\{\mu^i\}_{\mathbb{I}_1}, \{\mu^i_{s'}\}_{\mathbb{I}_2\times\mathbb{S}}\} \in \mathbb{R}^N$, where $N = |\mathbb{I}_1| + S|\mathbb{I}_2|$. The perturbed incentive problem (11) augments the objective from the original problem (10) with weighted sums of private evaluations $U^i(s',y^\infty)$. For agents $i \in \mathbb{I}_2$, these weights are allowed to depend on the current shock s'. Assumptions 3 and 4 ensure that Ω_1 is non-empty and that the objective in (11) is well defined and real-valued on this set. Hence, $V^* : \mathbb{S} \times \mathbb{R}^N \to \mathbb{R} \cup \{\infty\}$ is also well defined, though possibly infinite-valued. Let Dom $V^* = \{(s,\mu) \in \mathbb{S} \times \mathbb{R}^N \mid V^*(s,\mu) < \infty\}$ denote the *effective domain* of V^* on which V^* is finite. Anticipating our subsequent recursive formulation, we call $\mathbb{S} \times \mathbb{R}^N$ the *state space*.[8]

Proposition 1 gives properties of V^* that are easily found without recourse to recursive methods. We need the following definition. A function $g : \mathbb{S} \times \mathbb{R}^N \to \overline{\mathbb{R}}$ is inf-proper if for each $s \in \mathbb{S}$, $g(s,\cdot)$ is everywhere greater than $-\infty$ and somewhere less than ∞.

Proposition 1. *Let Assumptions 3 and 4 hold. i) V^* is inf-proper. ii) V^* is convex and, hence, each $V^*(s)$ is continuous and sub-differentiable on int Dom $V^*(s)$, where Dom $V^*(s) = \{\mu \in \mathbb{R}^N | V^*(s,\mu) < \infty\}$.*

Proof. See Appendix A. ☐

In particular, if Dom $V^* = \mathbb{S} \times \mathbb{R}^N$, then V^* is continuous.

[8]In some problems, such as Example 2, the optimal value is unaffected by the prior shock s and this argument may be dropped. Then $V^* : \mathbb{R}^N \to \mathbb{R} \cup \{\infty\}$ and Dom $V^* \subset \mathbb{R}^N$.

4.2 Deriving the functional equation

The first step in the application of the dual recursive approach is the formation of a Lagrangian:

$$\mathscr{L}_s(y^\infty,\eta;\mu) = F(s,y^\infty) + \sum_{\mathbb{I}_1} \mu^i \sum_{\mathbb{S}} U^i(s',y^\infty)Q(s,s') \\ + \sum_{\mathbb{I}_2} \sum_{\mathbb{S}} \mu^i_{s'} U^i(s',y^\infty)Q(s,s') + \sum_{\mathbb{M}} \eta^m G_m(y^\infty), \tag{12}$$

where $\eta = \{\eta^m\} \in \mathbb{R}^M_+$ is a non-negative multiplier. The Lagrangian incorporates *only* the current (date 1) incentive constraints. The perturbed incentive problem can be rewritten in primal sup-inf form as:

$$V^*(s,\mu) = \sup_{\Omega_2} \inf_{\mathbb{R}^M_+} \mathscr{L}_s(y^\infty,\eta;\mu). \tag{13}$$

Associated with the primal problem (13) is the *dual problem*:

$$V^D(s,\mu) = \inf_{\mathbb{R}^M_+} \sup_{\Omega_2} \mathscr{L}_s(y^\infty,\eta;\mu). \tag{14}$$

The dual (14) permits a very convenient decomposition of the inner supremum operation into a family of current and continuation problems linked by multipliers. Let $\mathscr{F}$ denote the set of inf-proper functions with domain $\mathbb{S} \times \mathbb{R}^N$. The (dual) Bellman operator $\mathscr{D}$ is defined next.

Definition 2. The Bellman operator $\mathscr{D}$ is defined on $\mathscr{F}$ as, for $V \in \mathscr{F}$ and each $(s,\mu) \in \mathbb{S} \times \mathbb{R}^N$,

$$\mathscr{D}(V)(s,\mu) = \inf_{\mathbb{R}^M_+} \sum_{\mathbb{S}} \left[\nu_{s,s'}(\mu,\eta) + \beta V\left(s', \mu'_{s,s'}(\mu,\eta)\right)\right] Q(s,s'), \tag{15}$$

with, for each $s' \in \mathbb{S}$,

$$\nu_{s,s'}(\mu,\eta) = \sup_{C(s')} f(s',y) + \sum_{\mathbb{M}} \frac{\eta^m}{Q(s,s')} u^m(s',y) + \sum_{\mathbb{I}_1} \mu^i u^i(s',y) + \sum_{\mathbb{I}_2} \mu^i_{s'} u^i(s',y), \tag{16}$$

and $\mu'_{s,s'} = \{\mu^{i\prime}_{s,s'}\}_{i \in \mathbb{I}}$ with

$$\mu^{i\prime}_{s,s'}(\mu, \eta) = \frac{\delta}{\beta}\Big[\mu^i + \sum_{\mathbb{M}} \frac{\eta^m q^{m,i}_{s'}}{Q(s,s')}\Big], \ i \in \mathbb{I}_1, \tag{17a}$$

$$\mu^{i\prime}_{s,s'}(\mu, \eta) = \left\{\frac{\delta}{\beta}\Big[\mu^i_{s'} + \sum_{\mathbb{M}} \frac{\eta^m q^{m,i}_{s',s''}}{Q(s,s')Q(s',s'')}\Big]\right\}_{s'' \in \mathbb{S}}, \ i \in \mathbb{I}_2. \tag{17b}$$

Proposition 2. *V^D satisfies: $V^D = \mathscr{D}(V^*)$.*

Proof. See Appendix A. □

For V^* to be a fixed point of $\mathscr{D}$, it is necessary that $V^* = V^D$. Standard weak duality results imply that $V^D \geq V^*$. However, under Assumptions 3 and 4 equality of value functions occurs and the minimum in $\mathscr{D}(V^*)(s,\mu)$ is attained whenever $V^*(s,\mu)$ is finite. Combining this (well known) result with Proposition 2 gives the following.

Proposition 3. *If Assumptions 3 and 4 hold and either Dom $V^* \subseteq \mathbb{R}^N_+$ or each function $u^i(s,\cdot)$ is affine, then V^* satisfies the Bellman equation:*

$$V^* = \mathscr{D}(V^*). \tag{18}$$

In addition, if $V^*(s,\mu) < \infty$, then:

$$\Gamma(s,\mu) := \operatorname*{argmin}_{\mathbb{R}^M_+} \sum_{\mathbb{S}} \left[v_{s,s'}(\mu,\eta) + \beta V^*\left(s', \mu'_{s,s'}(\mu,\eta)\right)\right] Q(s,s'),$$

is non-empty and coincides with the set of minimizers for (14).

Proof. See Appendix. □

In some problems, there is a set $A \subseteq \mathbb{R}^N_+$ such that for all $\eta \in \mathbb{R}^M_+$, $s, s' \in \mathbb{S}$, $\mu'_{s,s'}(\cdot,\eta) : A \to A$. For example, in the one sided limited commitment problem all constraint weights satisfy $q^m_{s,s'} = q^m_s Q(s,s') \geq 0$, with q^m_s equal to 1 if $m = s$ and 0 otherwise. Hence, for all $\eta \in \mathbb{R}^M_+$, $s, s' \in \mathbb{S}$, $\mu'_{s,s'}(\cdot,\eta) : \mathbb{R}^N_+ \to \mathbb{R}^N_+$. For problems of this sort, under Assumptions 3 and 4, $\forall (s,\mu) \in A$, $V^*(s,\mu) = \mathscr{D}(V^*)(s,\mu)$.

5 Contraction

This section contains the main result of the paper: it establishes sufficient conditions for $\mathscr{D}$ to be contractive on an appropriate space of functions. It adapts an argument of Rincon-Zapatero and Rodriguez-Palermo (2003). We restrict attention to problems in which Dom $V^* = \mathbb{S} \times \mathbb{R}^N$. This restriction includes many, though not all, examples (see Section 6) and, in particular, includes the one sided limited commitment problem. For a generalization of the arguments given below that covers other cases, see Appendix B. We make use of the following result from Messner, Pavoni, and Sleet (2011b).

Proposition 4. *Assume $V : \mathbb{S} \times \mathbb{R}^N \to \mathbb{R}$ is convex, then $\mathscr{D}(V)$ is also convex. In addition, if there exists a pair of functions $\underline{V}, \overline{V} : \mathbb{S} \times \mathbb{R}^N \to \mathbb{R}$ such that $\underline{V} \le \mathscr{D}(V) \le \overline{V}$, then $\mathscr{V}$ is real-valued and, hence, continuous.*

Proof. Follows from the proof of Lemma 4 in Messner, Pavoni, and Sleet (2011b). □

The key additional assumption needed to ensure contractivity is the following.

Assumption 5. *There exists a triple of functions: $\underline{\underline{V}} : \mathbb{S} \times \mathbb{R}^N \to \mathbb{R}$, $\underline{V} : \mathbb{S} \times \mathbb{R}^N \to \mathbb{R}$ and $\overline{V} : \mathbb{S} \times \mathbb{R}^N \to \mathbb{R}$ such that: (i) $\underline{\underline{V}} < \underline{V} \le V^* \le \overline{V}$, (ii)*

$$\inf_{\mathbb{S}\times\mathbb{R}^N} \underline{V}(s,\mu) - \underline{\underline{V}}(s,\mu) > 0 \tag{19a}$$

$$\sup_{\mathbb{S}\times\mathbb{R}^N} \frac{\overline{V}(s,\mu) - \underline{\underline{V}}(s,\mu)}{\underline{V}(s,\mu) - \underline{\underline{V}}(s,\mu)} = 1 + \sup_{\mathbb{S}\times\mathbb{R}^N} \frac{\overline{V}(s,\mu) - \underline{V}(s,\mu)}{\underline{V}(s,\mu) - \underline{\underline{V}}(s,\mu)} < \infty, \tag{19b}$$

(iii) $\mathscr{D}(\overline{V}) \le \overline{V}$ and (iv) $\underline{\underline{V}} \le \mathscr{D}(\underline{\underline{V}})$, $\underline{V} \le \mathscr{D}(\underline{V})$.

Notice that if $\underline{V} : \mathbb{S} \times \mathbb{R}^N \to \mathbb{R}$ and $\overline{V} : \mathbb{S} \times \mathbb{R}^N \to \mathbb{R}$ satisfy $\underline{V} \le \overline{V}$, then defining $\underline{\underline{V}} = -1 + 2\underline{V} - \overline{V}$ ensures the bounds in (19) hold since in this case:

$$\inf_{\mathbb{S}\times\mathbb{R}^N} \underline{V}(s,\mu) - \underline{\underline{V}}(s,\mu) = \inf_{\mathbb{S}\times\mathbb{R}^N} 1 + \overline{V}(s,\mu) - \underline{V}(s,\mu) > 0$$

$$1 + \sup_{\mathbb{S}\times\mathbb{R}^N} \frac{\overline{V}(s,\mu) - \underline{V}(s,\mu)}{\underline{V}(s,\mu) - \underline{\underline{V}}(s,\mu)} = 1 + \sup_{\mathbb{S}\times\mathbb{R}^N} \frac{\overline{V}(s,\mu) - \underline{V}(s,\mu)}{1 + \overline{V}(s,\mu) - \underline{V}(s,\mu)} < \infty.$$

Hence, for Assumption 5 to hold, it is sufficient that the following condition holds.

Assumption 6. *There exists a pair of functions $\underline{V} : \mathbb{S} \times \mathbb{R}^N \to \mathbb{R}$ and $\overline{V} : \mathbb{S} \times \mathbb{R}^N \to \mathbb{R}$ such that: (i) $\underline{V} \le V^* \le \overline{V}$, (ii) $\mathscr{D}(\overline{V}) \le \overline{V}$ and (iii) for all functions $V : \mathbb{S} \times \mathbb{R}^N \to \mathbb{R}$ with $V \le \underline{V}$, $V \le \mathscr{D}(V)$.*

Given a pair of functions $\underline{V}$ and $\overline{V}$ satisfying Assumption 5 (or Assumption 6), define the interval of functions:

$$\mathscr{V} = \{V : \mathbb{S} \times \mathbb{R}^N \to \mathbb{R} | V \text{ is convex (and, hence, continuous) and } \underline{V} \le V \le \overline{V}\}.$$

Define $d : \mathscr{V} \times \mathscr{V} \to \mathbb{R}_+$ according to:

$$d(V_1, V_2) = \sup_{\mathbb{S} \times \mathbb{R}^N} \left| \ln \left(\frac{V_1(s,\mu) - \underline{\underline{V}}(s,\mu)}{\underline{V}(s,\mu) - \underline{\underline{V}}(s,\mu)} \right) - \ln \left(\frac{V_2(s,\mu) - \underline{\underline{V}}(s,\mu)}{\underline{V}(s,\mu) - \underline{\underline{V}}(s,\mu)} \right) \right|$$
$$\leq \sup_{\mathbb{S} \times \mathbb{R}^N} \ln \left(\frac{\overline{V}(s,\mu) - \underline{\underline{V}}(s,\mu)}{\underline{V}(s,\mu) - \underline{\underline{V}}(s,\mu)} \right) < \infty,$$

where the finiteness stems from Assumption 5.

Lemma 1. *$(\mathscr{V}, d)$ is a complete metric space.*

Proof. See Appendix A. □

Proposition 5 verifies that $\mathscr{D}$ is contraction on $\mathscr{V}$.

Proposition 5. *Let Assumption 5 hold. There is an $\alpha \in [0,1)$ such that for all $V_1, V_2 \in \mathscr{V}$, $d(\mathscr{D}(V_1), \mathscr{D}(V_2)) \leq \alpha d(V_1, V_2)$.*

Proof. Let $V_1, V_2 \in \mathscr{V}$ and let $\lambda \in [0,1]$. Then, it is easily shown that: $\mathscr{D}(\lambda V_1 + (1-\lambda)V_2) \geq \lambda \mathscr{D}(V_1) + (1-\lambda)\mathscr{D}(V_2)$, i.e. $\mathscr{D}$ is concave on $\mathscr{V}$. By definition of d,

$$\ln \left(\frac{V_2 - \underline{\underline{V}}}{\underline{V} - \underline{\underline{V}}} \right) \leq \ln \left(\frac{V_1 - \underline{\underline{V}}}{\underline{V} - \underline{\underline{V}}} \right) + d(V_1, V_2).$$

Taking the exponential of each side at each (s, μ) and rearranging gives:

$$\exp\{-d(V_1, V_2)\} \left(\frac{V_2 - \underline{\underline{V}}}{\underline{V} - \underline{\underline{V}}} \right) \leq \left(\frac{V_1 - \underline{\underline{V}}}{\underline{V} - \underline{\underline{V}}} \right).$$

But $\underline{V} - \underline{\underline{V}} > 0$ and so: $\exp\{-d(V_1, V_2)\}(V_2 - \underline{\underline{V}}) \leq (V_1 - \underline{\underline{V}})$. Rearranging gives:

$$V_1 \geq \exp\{-d(V_1, V_2)\} V_2 + (1 - \exp\{-d(V_1, V_2)\}) \underline{\underline{V}}. \tag{20}$$

By monotonicity and concavity of $\mathscr{D}$:

$$\mathscr{D}(V_1) \geq \mathscr{D}(\exp\{-d(V_1, V_2)\} V_2 + (1 - \exp\{-d(V_1, V_2)\}) \underline{\underline{V}})$$
$$\geq \exp\{-d(V_1, V_2)\} \mathscr{D}(V_2) + (1 - \exp\{-d(V_1, V_2)\}) \mathscr{D}(\underline{\underline{V}}). \tag{21}$$

Equation (20) applied to $\mathscr{D}(\underline{V})$ and $\mathscr{D}(V_2)$ gives:

$$\mathscr{D}(\underline{V}) \geq \exp\{-d(\mathscr{D}(V_2), \mathscr{D}(\underline{V}))\} \mathscr{D}(V_2) + (1 - \exp\{-d(\mathscr{D}(V_2), \mathscr{D}(\underline{V}))\}) \underline{\underline{V}}. \tag{22}$$

Combining (21) with (22) gives:

$$\begin{aligned}\mathscr{D}(V_1) \geq{} & \exp\{-d(V_1,V_2)\}\mathscr{D}(V_2) + (1-\exp\{-d(V_1,V_2)\}) \\ & \times [\exp\{-d(\mathscr{D}(V_2),\mathscr{D}(\underline{\underline{V}}))\}\mathscr{D}(V_2) + (1-\exp\{-d(\mathscr{D}(V_2),\mathscr{D}(\underline{\underline{V}}))\})\underline{\underline{V}}].\end{aligned}$$

Letting $\mu = \exp\{-d(V_1,V_2)\} + (1-\exp\{-d(V_1,V_2)\})\exp\{-d(\mathscr{D}(V_2),\mathscr{D}(\underline{\underline{V}}))\}$, then gives:

$$\frac{\mathscr{D}(V_1)-\underline{\underline{V}}}{\underline{V}-\underline{\underline{V}}} \geq \mu\frac{\mathscr{D}(V_2)-\underline{\underline{V}}}{\underline{V}-\underline{\underline{V}}}.$$

Taking logs gives:

$$\ln\left(\frac{\mathscr{D}(V_1)-\underline{\underline{V}}}{\underline{V}-\underline{\underline{V}}}\right) \geq \ln\mu + \ln\left(\frac{\mathscr{D}(V_2)-\underline{\underline{V}}}{\underline{V}-\underline{\underline{V}}}\right).$$

By the definition of μ and Jensen's inequality:

$$\begin{aligned}\ln\mu &\geq (1-\exp\{-d(\mathscr{D}(V_2),\mathscr{D}(\underline{\underline{V}}))\})\ln\exp\{-d(V_1,V_2)\} + \exp\{-d(\mathscr{D}(V_2),\mathscr{D}(\underline{\underline{V}}))\})\ln 1 \\ &= -(1-\exp\{-d(\mathscr{D}(V_2),\mathscr{D}(\underline{\underline{V}}))\})d(V_1,V_2).\end{aligned}$$

Since $\underline{\underline{V}} \leq \mathscr{D}(\underline{\underline{V}}) \leq \mathscr{D}(\underline{V}) \leq \mathscr{D}(V_2) \leq \overline{V}$, it follows that $d(\mathscr{D}(V_2),\mathscr{D}(\underline{\underline{V}})) \leq d(\overline{V},\underline{\underline{V}}) < \infty$. Hence, $(1-\exp\{-d(\mathscr{D}(V_2),\mathscr{D}(\underline{\underline{V}}))\}) := \alpha \leq (1-\exp\{-d(\overline{V},\underline{\underline{V}})\}) < 1$. Thus,

$$\alpha d(V_1,V_2) \geq -\ln\mu \geq \ln\left(\frac{\mathscr{D}(V_2)-\underline{\underline{V}}}{\underline{V}-\underline{\underline{V}}}\right) - \ln\left(\frac{\mathscr{D}(V_1)-\underline{\underline{V}}}{\underline{V}-\underline{\underline{V}}}\right). \tag{23}$$

The same bound clearly holds with V_1 and V_2 on the right hand side of (23) interchanged. Hence, $\alpha d(V_1,V_2) \geq d(\mathscr{D}(V_1),\mathscr{D}(V_2))$. □

Application of the contraction mapping theorem then immediately yields Theorem 1, the main result of the paper.

Theorem 1. *Let Assumptions 3 to 5 hold. V^* is the unique fixed point of $\mathscr{D}$ in $\mathscr{V}$. Also, there is an $\alpha \in [0,1)$ such that for any $V \in \mathscr{V}$, $\mathscr{D}^n(V) \xrightarrow{d} V^*$ with $d(\mathscr{D}^n(V),V^*) \leq \alpha^n d(V,V^*) \leq \alpha^n d(\overline{V},\underline{V})$.*

Discussion of assumptions Assumptions 3 and 4 are basic assumptions that allow the application of duality results and are used to establish $V^* = \mathscr{D}(V^*)$. Assumption 5 is used to establish contractivity. The key difficulty in applying Theorem 1 involves finding bounding functions such that Assumption 5 holds. The construction of these will often be problem specific.

6 Applications

In this section we give various applications of Theorem 1.

Example 1 (One Sided Limited Commitment)**.** We first revisit the one sided limited commitment example from Section 2. First note that the earlier Assumptions 1 and 2 ensure that Assumptions 3 and 4 hold in this problem. Hence, Proposition 3 confirms that the principal's value function satisfies the Bellman equation:

$$V^*(s,\mu) = \mathscr{D}(V^*)(s,\mu) = \inf_{\mathbb{R}^M_+} \sum_{\mathbb{S}} \left[v_{s'}(\mu,\eta) + \beta V^* \left(s', \mu'_{s'}(\mu,\eta) \right) \right] Q(s,s'), \tag{24}$$

where, from (7) and Definition 2 or as derived previously in Section 2, $v_{s'}(\mu,\eta) = \sup_{y \in C(s')} (\mu + \eta^{s'})y + f(s',y)$ and $\mu'_{s'}(\mu,\eta) = \mu + \eta^{s'}$, with $C(s) = \{y : \omega(s) - \kappa(y + b(s)) \geq 0\}$, $f(s,y) = v(\omega(s) - \kappa(y + b(s)))$ and $\kappa = v^{-1}$.

Note that Assumptions 1 and 2, by ensuring the non-empty and compact-valuedness of C and the continuity of f, guarantee that Dom $V^* = \mathbb{S} \times \mathbb{R}$. A lower value function may be constructed from the more constrained problem in which y^∞ is restricted to equal 0, i.e. $\forall t, s^t, y_t(s^t) = 0 \in C(s_t)$. Then:

$$\underline{V}(s,\mu) = F(s,0).$$

Evidently, $\underline{V} : \mathbb{S} \times \mathbb{R} \to \mathbb{R}$. In addition, $\underline{V}$ is convex and continuous (in fact constant) in μ. Define the less constrained problem by:

$$\overline{V}(s,\mu) = \sup_{\Omega_0} F(s,y^\infty) + \mu \sum_{\mathbb{S}} U(s',y^\infty) Q(s,s').$$

It is easily checked that each $\overline{V}$ is convex, continuous and real-valued. In addition, it is easily verified that $\underline{V}$, $\overline{V}$ and any $V : \mathbb{S} \times \mathbb{R} \to \mathbb{R}$ with $V \leq \underline{V}$ satisfy: $\underline{V} \leq \mathscr{D}(\underline{V}) \leq V^* \leq \mathscr{D}(\overline{V}) \leq \overline{V}$ and $V \leq \mathscr{D}(V)$. Hence, Assumption 6 holds and Theorem 1 is applicable.

Proposition 6. *Let Assumption 1 hold and let $\mathscr{V}$ denote the interval of convex (and continuous) functions bounded below by $\underline{V}$ and above by $\overline{V}$. V^* is the unique fixed point of $\mathscr{D}$ in $\mathscr{V}$. Also, there is an $\alpha \in [0,1)$ such that for any $V \in \mathscr{V}$, $\mathscr{D}^n(V) \to V^*$ with $d(\mathscr{D}^n(V), V^*) \leq \alpha^n d(V,V^*) \leq \alpha^n d(\overline{V}, \underline{V})$.*

Proof. Follows directly from Theorem 1. ☐

This example is easily extended to one with a committed principal and multiple uncommitted agents or, following the treatment in Messner, Pavoni, and Sleet (2011b), one with no committed principal and only uncommitted agents.

Example 2 (Hidden information: IID shocks)**.** A principal insures an agent against privately observed taste shocks. The agent's taste shocks $\{s_t\}_{t=1}^{\infty}$ take values in the finite set $\mathbb{S} = \{1, \ldots, S\}$ and are i.i.d. with per period distribution $Q \in \mathbb{R}^S$. The agent values lifetime consumption plans $c^\infty = \{c_t\}_{t=1}^{\infty}$, $c_t : \mathbb{S}^t \to \mathbb{R}_+$, according to:

$$\sum_{t=1}^{\infty} \delta^{t-1} E[s_t v(c_t(s^t))], \tag{25}$$

where $\delta \in (0,1)$ and $v : \mathbb{R}_+ \to Y$ satisfies the following condition.

Assumption 7. *$v : \mathbb{R}_+ \to Y$ is increasing, strictly concave and bounded: $Y = [\underline{y}, \overline{y})$, $-\infty < \underline{y} < \overline{y} < \infty$.*

Given a consumption plan c^∞, the agent's corresponding utility plan is $y^\infty = \{y_t\}$, where $y_t(s^t) = v(c_t(s^t))$. It is convenient to re-express the analysis in terms of utility plans. Let

$$U(s, y^\infty) = \sum_{t=1}^{\infty} \delta^{t-1} E[s_t y_t(s^t) | s_1 = s'] \tag{26}$$

denote the agent's lifetime payoff from utility plan y^∞ conditional on the first period shock s'. The principal's per period objective is $f(s,y) = sy - \psi\kappa(y)$, where $\psi \in \mathbb{R}_{++}$ is the shadow price of resources and $\kappa : Y \to \mathbb{R}_+$ is the inverse of v. The principal uses a discount factor $\beta \in [0,1)$ and has lifetime payoff is: $F(y^\infty) = \sum_{t=1}^{\infty} \beta^{t-1} E[f(s_t, y_t(s^t))]$. Since f is bounded above, F is well defined. Atkeson and Lucas (1992) assume $\delta = \beta < 1$; Farhi and Werning (2007) assume $\delta < \beta < 1$. Let $\Omega_0 = \{y^\infty | \forall t, y_t : \mathbb{S}^t \to Y \text{ and } F(y^\infty) > -\infty\}$ denote the set of plans with finite planner payoffs.

Without loss of generality attention may be restricted to utility plans that induce the agent to truthfully report the shocks she receives. Such plans satisfy the incentive-compatibility conditions, for all t, s^{t-1}, $s' \neq \hat{s}$,

$$\begin{aligned} s' y_t(s^{t-1}, s') + \delta \sum_{s' \in \mathbb{S}} U(s'', y_{t+1}^{\infty}(s^{t-1}, s')) Q(s'') \geq \\ s' y_t(s^{t-1}, \hat{s}) + \delta \sum_{s'' \in \mathbb{S}} U(s'', y_{t+1}^{\infty}(s^{t-1}, \hat{s})) Q(s''), \end{aligned} \tag{27}$$

where s' denotes a true shock, $\widehat{s}$ a lie and $y^{\infty}_{t+1}(s^t)$ is the continuation of y^{∞} after s^t. Let Ω_1 and Ω_2 denote the set of plans in Ω_0 satisfying these constraints from the first and second periods onwards respectively. Let $C(s) = Y$ for all s, $\mathbb{I} = \{1\}$ and $u^1(s,y) = sy$. Index constraints by true-lie shock pairs $(s', \widehat{s})$ and (after comparing (27) to the general form (6)) define:

$$u^{s,\widehat{s}}(\bar{s},y) = \begin{cases} sy & \text{if } \bar{s} = s \\ -sy & \text{if } \bar{s} = \widehat{s} \\ 0 & \text{otherwise} \end{cases} \quad \text{and} \quad q^{s,\widehat{s}}_{\bar{s},s'} = \begin{cases} Q(s') & \text{if } \bar{s} = s \\ -Q(s') & \text{if } \bar{s} = \widehat{s} \\ 0 & \text{otherwise.} \end{cases} \tag{28}$$

It is then easy to verify that Assumptions 3 and 4 follow from Assumption 7.

The principal's problem is: $\sup_{\Omega_1} F(y^{\infty})$. To solve this problem we embed it into the family of "perturbed" problems:

$$V^*(\mu) = \sup_{\Omega_1} F(y^{\infty}) + \mu \sum_{\mathbb{S}} U(s, y^{\infty}) Q(s), \tag{29}$$

where $V^* : \mathbb{R} \to \mathbb{R}$, and the original problem is recovered by setting $\mu = 0$. Using (28), Definition 2 and Proposition 3, we have for all $\mu \in \mathbb{R}$,

$$V^*(\mu) = \mathscr{D}(V^*)(\mu) = \inf_{\mathbb{R}^M_+} \sum_{\mathbb{S}} \{ \nu_{s'}(\mu,\eta) + \beta V^*(\mu'_{s'}(\mu,\eta)) \} Q(s'), \tag{30}$$

with

$$\nu_{s'}(\mu,\eta) = \sup_{[\underline{y},\overline{y})} \left(1 + \mu + \frac{1}{Q(s')} \left[\sum_{\widehat{s} \neq s'} \eta_{s',\widehat{s}} - \sum_{\widehat{s} \neq s'} \frac{\widehat{s}}{s'} \eta_{\widehat{s},s'} \right] \right) s'y - \psi\kappa(y), \tag{31a}$$

and

$$\mu'_{s'}(\mu,\eta) = \frac{\delta}{\beta} \left(1 + \mu + \frac{1}{Q(s')} \left[\sum_{\widehat{s} \neq s'} \eta_{s',\widehat{s}} - \sum_{\widehat{s} \neq s'} \eta_{\widehat{s},s'} \right] \right). \tag{31b}$$

$\nu_{s'}$ can be re-expressed as $\nu_{s'}(\mu,\eta) = \nu(\rho_{s'}(\mu,\eta))$, where:

$$\nu(\rho) = \sup_{Y} \rho y - \widetilde{\kappa}(y), \quad \widetilde{\kappa}(y) = \begin{cases} \psi\kappa(y) & \text{if } y \in Y \\ \infty & \text{otherwise,} \end{cases}$$

and $\rho_s(\mu,\eta) = (1 + \mu + \frac{1}{Q(s)}[\sum_{\widehat{s} \neq s} \eta_{s,\widehat{s}} - \sum_{\widehat{s} \neq s} \frac{\widehat{s}}{s} \eta_{\widehat{s},s}])s$. A function $f^{\star} : \mathbb{R}^n \to \overline{\mathbb{R}}$ is said to be the *conjugate* of $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ if for all $x^\star \in \mathbb{R}^n$, $f^\star(x^\star) = \sup_{\mathbb{R}^n}\langle x, x^\star\rangle - f(x)$, where $\langle\cdot,\cdot\rangle : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ gives the dot product of two elements of $\mathbb{R}^n$. Thus, ν is the conjugate of $\widetilde{\kappa}$. Properties of ν follow from those of κ and well known properties of conjugate functions, see Rockafellar (1970). For example, ν is real-valued, convex and, hence, continuous. The continuity of each $\nu_{s'}$ follows.

We define the following more and less constrained problems. The first is the *no insurance problem:*

$$\underline{V}(\mu) = \sup_{y^\infty \in \underline{\Omega}} F(y^\infty) + \mu \sum_{s' \in \mathbb{S}} U(s', y^\infty) Q(s'), \tag{32}$$

with $\underline{\Omega} = \{y^\infty | \exists y \in Y, \forall t, s^t, y_t(s^t) = y\}$; the second is the *full insurance problem:*

$$\overline{V}(\mu) = \sup_{y^\infty \in \overline{\Omega}} F(y^\infty) + \mu \sum_{s' \in \mathbb{S}} U(s', y^\infty) Q(s'), \tag{33}$$

with $\overline{\Omega} = \Omega_0$. It is easy to verify that the functions $\underline{V}$ and $\overline{V}$, satisfy:

$$\underline{V}(\mu) = \sum_{t=1}^{\infty} \beta^{t-1} \nu\Big(\Big(1 + \Big(\frac{\delta}{\beta}\Big)^{t-1} \mu\Big) E[s']\Big) \text{ and } \overline{V}(\mu) = \sum_{t=1}^{\infty} \beta^{t-1} E\Big[\nu\Big(\Big(1 + \Big(\frac{\delta}{\beta}\Big)^{t-1} \mu\Big) s'\Big)\Big].$$

Of course, $\underline{V} \leq V^* \leq \overline{V}$ and $\underline{V}, \overline{V} : \mathbb{S} \times \mathbb{R}^N \to \mathbb{R}$. It is easy to see that $\underline{V} \leq \mathscr{D}(\underline{V}) \leq V^* \leq \mathscr{D}(\overline{V}) \leq \overline{V}$. In addition, if $V \leq \underline{V}$, then $V \leq \mathscr{D}(V)$, thus, Assumption 6 holds. It also follows from the properties of ν that $\underline{V}$ and $\overline{V}$ are convex (and continuous). Thus, Assumption 6 is verified. The following result is then an immediate corollary of Theorem 1.

Proposition 7. *Let Assumption 7 hold and let $\mathscr{V}$ denote the interval of convex (and continuous) functions bounded below by $\underline{V}$ and above by $\overline{V}$. V^* is the unique fixed point of $\mathscr{D}$ in $\mathscr{V}$. Also, there is a $\alpha \in [0,1)$ such that for any $V \in \mathscr{V}$, $\mathscr{D}^n(V) \to V^*$ with $d(\mathscr{D}^n(V), V^*) \leq \alpha^n d(V, V^*) \leq \alpha^n d(\overline{V}, \underline{V})$.*

Example 3 (Hidden information: Markov shocks)**.** Let everything be as before except that now shocks evolve according to a Markov process with transition Q from a seed $s_0 = s$. The planner's problem is $\sup_{\Omega_1} F(s, y^\infty)$, where $F(s, y^\infty) = \sum_{t=1}^{\infty} \beta^{t-1} E[f(s_t, y_t^\infty(s^t))|s_0 = s]$ and Ω_1 is modified to include Markov shocks:

$$s' y_t(s^{t-1}, s') + \delta \sum_{s'' \in \mathbb{S}} U(s'', y_{t+1}^\infty(s^{t-1}, s')) Q(s', s'') \geq \\ s' y_t(s^{t-1}, \widehat{s}) + \delta \sum_{s'' \in \mathbb{S}} U(s'', y_{t+1}^\infty(s^{t-1}, \widehat{s})) Q(s', s''). \tag{34}$$

Once again Assumption 7 ensures Assumptions 3 and 4 hold. To derive a recursive formulation of the principal's problem it is necessary to expand the dimension of the weight state variable and use an ex post weight $\mu = \{\mu_{s'}\}_{s' \in \mathbb{S}} \in \mathbb{R}^S$. The perturbed problem becomes, $\forall (s, \mu) \in \mathbb{S} \times \mathbb{R}^S$,

$$V^*(s, \mu) = \sup_{\Omega_1} F(s, y^\infty) + \sum_{\mathbb{S}} \mu_{s'} U(s', y^\infty) Q(s, s'), \tag{35}$$

Using

$$u^{s,\widehat{s}}(\bar{s}, y) = \begin{cases} sy & \text{if } \bar{s} = s \\ -sy & \text{if } \bar{s} = \widehat{s} \\ 0 & \text{otherwise} \end{cases} \quad \text{and} \quad q^{s,\widehat{s}}_{\bar{s},s'} = \begin{cases} Q(s, s') & \text{if } \bar{s} = s \\ -Q(s, s') & \text{if } \bar{s} = \widehat{s} \\ 0 & \text{otherwise,} \end{cases} \tag{36}$$

and Definition 2 and Proposition 3, we have for all $(s, \mu) \in \mathbb{S} \times \mathbb{R}^S$,

$$\mathscr{D}^*(V)(s, \mu) = \inf_{\mathbb{R}^M_+} \sum_{\mathbb{S}} \left[\nu_{s,s'}(\mu, \eta) + \beta V^* \left(s', \mu'_{s,s'}(\mu, \eta) \right) \right] Q(s, s'),$$

and

$$\nu_{s,s'}(\mu, \eta) = \sup_{y \in Y} \left(1 + \mu_{s'} + \sum_{\widehat{s} \neq s'} \eta_{s',\widehat{s}} - \sum_{\widehat{s} \neq s'} \eta_{\widehat{s},s'} \frac{\widehat{s}}{s'} \frac{Q(s, \widehat{s})}{Q(s, s')} \right) s'y - \psi\kappa(y).$$

Weights are now updated according to $\mu_{s,s'} = \{\mu_{s,s',s''}\}_{s'' \in \mathbb{S}}$,

$$\mu'_{s,s',s''}(\mu, \eta) = \frac{\delta}{\beta} \left(\mu_{s'} + \sum_{\widehat{s} \neq s'} \eta_{s',\widehat{s}} - \sum_{\widehat{s} \neq s'} \eta_{\widehat{s},s'} \frac{Q(\widehat{s}, s'')}{Q(s', s'')} \frac{Q(s, \widehat{s})}{Q(s, s')} \right).$$

No and full insurance bounding value functions $\underline{V}, \overline{V} : \mathbb{S} \times \mathbb{R}^S \to \mathbb{R}$ can be defined analogously to the i.i.d. case and Assumption 6 verified. The following result is then an immediate corollary of Theorem 1.

Proposition 8. *Let Assumption 7 hold and let $\mathscr{V}$ denote the interval of convex (and continuous) functions bounded below by $\underline{V}$ and above by $\overline{V}$. V^* is the unique fixed point of $\mathscr{D}$ in $\mathscr{V}$. Also, there is a $\alpha \in [0, 1)$ such that for any $V \in \mathscr{V}$, $\mathscr{D}^n(V) \to V^*$ with $d(\mathscr{D}^n(V), V^*) \leq \alpha^n d(V, V^*) \leq \alpha^n d(\overline{V}, \underline{V})$.*

7 Policies

The application of $\mathscr{D}$ to the function V^* induces a policy correspondence Ψ where:

$$\Psi(s,\mu) = \{(\eta,\mu')|\eta \in \Gamma(s,\mu) \text{ and } \mu' = \{\mu_{s,s'}(\mu,\eta)\}_{s'}\}.$$

We say that $\{\eta_t, \mu_{t+1}\}$, $\eta_t : \mathbb{S}^{t-1} \to \mathbb{R}^M_+$ and $\mu_{t+1} : \mathbb{S}^t \to \mathbb{R}^M$, is generated by Ψ from (s_0,μ_1) if $(\eta_1,\mu_2) \in \Psi(s_0,\mu_1)$ and for each $s^{t-1} = (s^{t-2}, s_{t-1})$, $(\eta_t(s^{t-1}), \mu_{t+1}(s^{t-1})) \in \Psi(s_{t-1}, \mu_t(s^{t-1}))$. Let:

$$\Xi(s,\mu,\eta) = \left\{ y \in \prod_{s' \in \mathbb{S}} C(s') \;\middle|\; \begin{array}{r} \forall s', y_{s'} \in \text{argmax}_{C(s')} f(s',y) + \sum_{\mathbb{M}} \frac{\eta^m}{Q(s,s')} u^m(s',y) \\ + \sum_{\mathbb{I}_1} \mu^i u^i(s',y) + \sum_{\mathbb{I}_2} \mu^i_{s'} u^i(s',y) \end{array} \right\}.$$

We say that a plan $\{y_t\}$ is generated by (Ψ, Ξ) from (s_0, μ_1) if there is a sequence $\{\eta_t, \mu_{t+1}\}$ generated by Ψ such that each $y_t(s^{t-1}) \in \Xi(s_{t-1}, \mu_t(s^{t-1}), \eta_t(s^{t-1}))$. Messner, Pavoni, and Sleet (2011b), Proposition 7, shows that if $\{y^*_t\}$ solves the generalized incentive problem (11) at (s_0,μ_1) and if the Lagrangians $\mathscr{L}_s(\cdot,\cdot;\mu)$ admit saddle points at all $(s,\mu) \in$ Dom V^*, then $\{y^*_t\}$ solves the recursive dual problem in the sense that is generated by (Ψ,Ξ) from (s_0,μ_1). However, even under these conditions (Ψ,Ξ) may generate extraneous action-plans that are not optimal for (11). Thus, the recursive dual gives necessary, but not sufficient conditions for a primal optimum.[9] This difficulty is resolved if (11) admits a solution and (Ψ,Ξ) generates a unique plan (from some (s_0,μ_1)). Consequently, the imposition of additional strict concavity conditions ensures that the recursive dual gives sufficient conditions for a primal optimum, see Messner, Pavoni, and Sleet (2011b), Proposition 8. [10]

7.1 Feasible weight processes and the long run behavior of plans

Even if the recursive dual problem gives only necessary conditions for an optimum some useful characterization of the optimum may be available. We conclude this paper with some brief remarks in this direction. The set of *all* non-negative valued multiplier processes (not necessarily optimal ones) and the updating functions (17a) and (17b) define a family of feasible weight processes.[11] The structure of the updating functions in a given

[9] As was pointed out by Messner and Pavoni (2004), the same issue arises in the context of Marcet and Marimon (2011)'s recursive saddle point method.

[10] Marcet and Marimon (2011) provide a related result for their recursive saddle point method. Cole and Kubler (2010) provide an extended recursive saddle point method that resolves the problem in weakly concave settings.

[11] i.e. for each $\{\eta_t\}$, $\eta_t : S^{t-1} \to \mathbb{R}^M_+$, $s_0 \in S$, $\mu_1 \in$ Dom V^* and $\mu_{t+1}(s^t) = \mu'_{s_t,s_{t+1}}(\mu_t(s^{t-1}), \eta_t(s^{t-1}))$.

problem often places useful restrictions on *all* such processes. In certain applications, these functions directly imply that all feasible weight processes (including the optimal one) are martingales, sub-martingales or a particular type of Markov process called an iterated function system.[12] Establishing these results plays a key role in characterizing the long run properties of the action plan induced by a feasible (and, hence, an optimal) process

For simplicity, consider the updating formula (17a) with decomposable weights and a single agent $i \in \mathbb{I}_1$. Dropping the agent index i from the notation, note that for any $\eta \in \mathbb{R}^M_+$,

$$\sum_{s' \in \mathbb{S}} \mu'_{s,s'}(\mu, \eta) Q(s, s') = \frac{\delta}{\beta}\Big[\mu + \sum_{m \in \mathbb{M}} \eta^m \sum_{s' \in \mathbb{S}} q^m_{s'}\Big]. \tag{37}$$

In the models of Atkeson and Lucas (1992) and Farhi and Werning (2007) and many other hidden information models with i.i.d. shocks, the terms $\sum_{s' \in \mathbb{S}} q^m_{s'}$ equal 0. Thus, for these models, (37) reduces to:

$$\sum_{s' \in \mathbb{S}} \mu'_{s'}(\mu, \eta) Q(s') = \frac{\delta}{\beta}\mu. \tag{38}$$

In particular, if the principal and agent discount at the same rate, $\beta = \delta$, then the weight process is *constrained to be* a martingale. If, in addition, there is a set $A \subseteq \text{Dom } V^* \cap \mathbb{R}_+$ such that at each $\mu \in A$ there is an optimal choice of η with $\mu'_s(\mu, \eta) \in A$, then the associated optimal weight process is an L_1-bounded martingale and the martingale convergence theorem implies that it almost surely converges. If the correspondence $\Xi(s, \mu, \eta)$ is single-valued (on Dom V^* and for such processes), then any action plan that is optimal given the multipliers must converges as well.[13] If $\delta < \beta$, as in Farhi-Werning, weight processes are constrained to belong to a class of Markov processes induced by an iterated function system. Limit theorems for such Markov processes can be used to show that, under certain circumstances, all weight processes (and, hence, the any optimal action plan) is ergodic.[14]

[12]An iterated function system is defined by a finite shock set $\mathbb{S}$, a probability distribution Q on $\mathbb{S}$, a state space X and a function $g : \mathbb{S} \times X \to X$. This tuple is used to defines a stochastic process satisfying: $x_{t+1} = g(s_t, x_t)$. Often each $g(s, \cdot)$ is assumed to be Lipschitz and $\sum_{\mathbb{S}} g(s, \cdot)Q(s)$ contractive.

[13]If an optimal multiplier function η^* can be shown to be non-zero at each interior point of Dom V^*, then convergence must occur to the boundary of Dom V^*. This is the basis of various "immiseration" results in the literature which assert almost sure convergence of μ to 0 and, hence, of agent utility to its lowest bound.

[14]We do not pursue these issues further here. See Sleet and Yeltekin (2009) for analysis of an optimal contracting problem that exploits this feature of the optimal weight process.

For the one-sided commitment model $\sum_{s'} q^m_{s'} = 1$ and

$$\sum_{s' \in \mathbb{S}} \mu'_{s,s'}(\mu, \eta) Q(s, s') = \frac{\delta}{\beta}\Big[\mu + \sum_{m \in \mathbb{M}} \eta^m\Big] \geq \frac{\delta}{\beta}\mu^i. \tag{39}$$

In this case all scaled weight processes $\{(\beta/\delta)^t \mu^i_t\}$ are constrained to be sub-martingales.

For the Atkeson-Lucas or Farhi-Werning model with persistent shocks, $\sum_{s',s''} q^m_{s',s''} = 0$. In these cases, the updating functions imply:

$$\sum_{s' \in \mathbb{S}} \sum_{s'' \in \mathbb{S}} \mu^{i\prime}_{s,s',s''}(\mu, \eta) Q(s', s'') Q(s, s') = \frac{\delta}{\beta} \sum_{s' \in \mathbb{S}} \mu^i_{s'} Q(s, s'). \tag{40}$$

Consequently, when $\beta = \delta$, the conditional expectation of the weight process $E_{t-1}[\mu_{t+1}]$ is constrained to be a martingale. When $\delta < \beta$, it is an iterated function system.

Appendices

A Proofs

Proof of Proposition 1. By Assumptions 3 and 4, for each $s \in \mathbb{S}$, Ω_1 contains a plan such that $F(s, \cdot)$ and each $U^i(s', \cdot)$ are real-valued. Hence, $V^*(s, \mu) > -\infty$. Also, $V^*(s, 0) = \sup_{\Omega_1} F(s, y^\infty) < \infty$ since f is bounded above. Thus, V^* is inf-proper. For $s \in \mathbb{S}$, let $\langle \mu, U \rangle = \sum_{\mathbb{I}_1} \mu^i \sum_{\mathbb{S}} U^i(s', y^\infty) Q(s, s') + \sum_{\mathbb{I}_2} \sum_{\mathbb{S}} \mu^i_{s'} U^i(s', y^\infty) Q(s, s')$. For $\mu(j) \in \mathbb{R}^N$, $j = 1, 2$ and $\theta \in [0, 1]$, let $\mu(\theta) = \theta\mu(1) + (1-\theta)\mu(2)$. We have:

$$\begin{aligned}
V^*(s, \mu(\theta)) &= \sup_{\Omega_1} F(s, y^\infty) + \langle \mu(\theta), U \rangle \\
&= \sup_{\Omega_1} F(s, y^\infty) + \theta \langle \mu(1), U \rangle + (1-\theta)\langle \mu(2), U \rangle \\
&\leq \theta \sup_{\Omega_1} \{F(s, y^\infty) + \langle \mu(1), U \rangle\} + (1-\theta) \sup_{\Omega_1} \{F(s, y^\infty) + \langle \mu(2), U \rangle\} \\
&= \theta V^*(s, \mu(1)) + (1-\theta) V^*(s, \mu(2)). \quad \square
\end{aligned}$$

Hence, V^* is convex. □

Proof of Proposition 2. Rearranging the terms in the Lagrangian yields the following specification of the dual:

$$V^D(s,\mu) = \inf_{\mathbb{R}_+^M} \sup_{\Omega_2} \sum_{\mathbb{S}} \Big\{ f(s',y(s')) + \sum_{\mathbb{M}} \frac{\eta^m}{Q(s,s')} u^m(s',y(s')) + \sum_{\mathbb{I}_1} \mu^i u^i(s',y(s')) + \sum_{\mathbb{I}_2} \mu^i_{s'} u^i(s',y(s'))$$
$$+ \beta \Big(F(s',y_2^\infty(s')) + \frac{\delta}{\beta} \sum_{\mathbb{I}_1} \Big\{ \mu^i + \sum_{\mathbb{M}} \frac{\eta^m q^{m,i}_{s'}}{Q(s,s')} \Big\} \sum_{\mathbb{S}} U^i(s'',y_2^\infty(s')) Q(s',s'')$$
$$+ \Big[\frac{\delta}{\beta} \sum_{\mathbb{I}_2} \sum_{\mathbb{S}} \Big\{ \mu^i_{s'} + \sum_{\mathbb{M}} \frac{\eta^m q^{m,i}_{s',s''}}{Q(s,s')Q(s',s'')} \Big\} U^i(s'',y_2^\infty(s')) \Big] Q(s',s'') \Big) \Big\} Q(s,s'). \tag{41}$$

The inner supremum operation of (41) can be decomposed into S current and S continuation maximizations. Using the definition of $\mu'_{s,s'}$ (with $\mu^{i'}_{s,s',s''}$ denoting the s'' element of $\mu^{i'}_{s,s'}$, $i \in \mathbb{I}_2$) we have:

$$V^D(s,\mu) = \inf_{\mathbb{R}_+^M} \sum_{\mathbb{S}} \Big[\sup_{C(s')} \Big\{ f(s',y) + \sum_{\mathbb{M}} \frac{\eta^m}{Q(s,s')} u^m(s,y) + \sum_{\mathbb{I}_1} \mu^i u^i(s',y) + \sum_{\mathbb{I}_2} \mu^i_{s'} u^i(s',y) \Big\}$$
$$+ \beta \sup_{\Omega_1} \Big\{ F(s',y^\infty) + \sum_{\mathbb{I}_1} \mu^{i'}_{s,s'}(\mu,\eta) \sum_{\mathbb{S}} U^i(s'',y^\infty) Q(s',s'')$$
$$+ \sum_{\mathbb{I}_2,\mathbb{S}} \mu^{i'}_{s,s',s''}(\mu,\eta) U^i(s'',y^\infty) Q(s',s'') \Big\} \Big] Q(s,s'),$$

The result then follows from the definitions of V^* and $\mathscr{D}$. □

Proof of Proposition 3. Given (s,μ), let $H(y^\infty) = F(s,y^\infty) + \sum_{\mathbb{I}_1} \mu^i \sum_{\mathbb{S}} U^i(s',y^\infty)Q(s,s') + \sum_{\mathbb{I}_2} \sum_{\mathbb{S}} \mu^i_{s'} U^i(s',y^\infty)Q(s,s')$. If $V^*(s,\mu) = \infty$, then for every $v \in \mathbb{R}$, there is a $y^{\infty,v} \in \Omega_1$ such that $H(y^{\infty,v}) \geq v$ and so, for all $\eta \in \mathbb{R}_+^M$, $\mathscr{L}_s(y^{\infty,v},\eta;\mu) = H(y^{\infty,v}) + \sum_{\mathbb{M}} \eta^m G_m(y^{\infty,v}) \geq v$. Thus, for all such v,

$$V^D(s,\mu) = \inf_{\mathbb{R}_+^M} \sup_{\Omega_1} \mathscr{L}_s(y^\infty,\eta;\mu) \geq \inf_{\mathbb{R}_+^M} \mathscr{L}_s(y^{\infty,v},\eta;\mu) \geq v.$$

It follows that $\mathscr{D}(V^*)(s,\mu) = V^D(s,\mu) = \infty = V^*(s,\mu)$. Next suppose that $V^*(s,\mu) \in \mathbb{R}$ (i.e. $(s,\mu) \in$ Dom V^*). By Assumptions 3 and 4 and the condition in the proposition, H is real-valued, concave, G is real-valued, concave and Ω_2 is convex. In addition, by Assumption 3, a Slater condition is satisfied. Thus, by Luenberger (1969), Theorem 1, p. 224, and Proposition 2, $V^*(s,\mu) = V^D(s,\mu) = \mathscr{D}(V^*)(s,\mu)$ and $\Gamma(s,\mu) = \operatorname{argmin}_{\mathbb{R}_+^M} \sup_{\Omega_2} \mathscr{L}_s(y^\infty,\eta;\mu) \neq \emptyset$. □

Proof of Lemma 1. Evidently, $(\mathscr{V},d)$ is a metric space. Let $\{V_n\}$ be a Cauchy sequence in

$\mathscr{V}$. Thus, as $n, m \to \infty$,

$$d(V_n, V_m) = \sup_{\mathbb{S}\times\mathbb{R}^N} \left| \ln\left(\frac{V_n(s,\mu) - \underline{\underline{V}}(s,\mu)}{\underline{V}(s,\mu) - \underline{\underline{V}}(s,\mu)}\right) - \ln\left(\frac{V_m(s,\mu) - \underline{\underline{V}}(s,\mu)}{\underline{V}(s,\mu) - \underline{\underline{V}}(s,\mu)}\right)\right| \to 0.$$

For each $n \in \mathbb{N}$, define $g_n : \mathbb{S} \times \mathbb{R}^N \to \mathbb{R}$ according to: $g_n(s,\mu) = \ln\left(\frac{V_n(s,\mu)-\underline{\underline{V}}(s,\mu)}{\underline{V}(s,\mu)-\underline{\underline{V}}(s,\mu)}\right)$, $(s,\mu) \in \mathbb{S} \times \mathbb{R}^N$. It follows that $\{g_n\}$ is Cauchy with respect to the sup-norm. Let $\mathscr{G} = \left\{g : \mathbb{S} \times \mathbb{R}^N \to \mathbb{R} \mid \exists V \in \mathscr{V} \text{ with } \forall (s,\mu), g(s,\mu) = \ln\left(\frac{V(s,\mu)-\underline{\underline{V}}(s,\mu)}{\underline{V}(s,\mu)-\underline{\underline{V}}(s,\mu)}\right)\right\}$. This is a closed subset of the Banach space of continuous, (sup-norm) bounded functions with domain $\mathbb{S} \times \mathbb{R}^N$. Each g_n belongs to $\mathscr{G}$ and since $\mathscr{G}$ is complete with respect to the sup norm, g_n converges with respect to the metric $\rho(g', g'') = \sup_{\mathbb{S}\times\mathbb{R}^N} |g'(s,\mu) - g''(s,\mu)|$ to a continuous, sup-norm bounded function g. Members of $\mathscr{G}$ are bounded below by 0 and above by $\ln\left(\frac{\overline{V}-\underline{\underline{V}}}{\underline{V}-\underline{\underline{V}}}\right)$ Let $V = \underline{\underline{V}} + \exp\{g\}(\underline{V} - \underline{\underline{V}})$. Evidently, $V_n \xrightarrow{d} V$ and, given the properties of g, $V \in \mathscr{V}$. $\square$

B An Extension

In this appendix the analysis is extended to the case in which Dom V^* is known a priori, but is a strict subset of $\mathbb{S} \times \mathbb{R}^N$. In this case, convexity of V^* does not ensure continuity. The following definitions will be used. A function $g : \mathbb{S} \times \mathbb{R}^N \to \mathbb{R} \cup \{\infty\}$ is (lower) level bounded if for each $s \in \mathbb{S}$ and $b \in \mathbb{R}$, the lower level set $\{y | g(s,y) \leq b\}$ is bounded. A function $g : \mathbb{S} \times \mathbb{R}^{N_1} \times \mathbb{R}^{N_2} \to \mathbb{R} \cup \{\infty\}$ is (lower) level bounded locally in its second argument if for each $s \in \mathbb{S}$, $\mu \in \mathbb{R}^{N_1}$ and $b \in \mathbb{R}$, there is a neighborhood $\mathscr{N}$ of μ and a bounded set $B \subset \mathbb{R}$ such that for each $\mu' \in \mathscr{N}$, $\{y | g(s,\mu',y) \leq b\} \subset B$.

Assumption 5 is modified slightly to give Assumption 5′ and supplemented with Assumption 8 below.

Assumption 5′. *There exists a triple of functions:* $\underline{\underline{V}} : \mathbb{S} \to \mathbb{R}$, $\underline{V} : \mathbb{S} \times \mathbb{R}^N \to \mathbb{R}$ *and* $\overline{V} : \mathbb{S} \times \mathbb{R}^N \to \mathbb{R}$ *such that:* $\underline{\underline{V}} < \underline{V} \leq V^* \leq \overline{V}$,

$$0 < \inf_{Dom\ V^*} \underline{V}(s,\mu) - \underline{\underline{V}}(s,\mu) \tag{42a}$$

$$\infty > \sup_{Dom\ V^*} \frac{\overline{V}(s,\mu) - \underline{\underline{V}}(s,\mu)}{\underline{V}(s,\mu) - \underline{\underline{V}}(s,\mu)} = 1 + \sup_{Dom\ V^*} \frac{\overline{V}(s,\mu) - \underline{V}(s,\mu)}{\underline{V}(s,\mu) - \underline{\underline{V}}(s,\mu)}, \tag{42b}$$

$\underline{\underline{V}} \leq \mathscr{D}(\underline{\underline{V}})$, $\underline{V} \leq \mathscr{D}(\underline{V})$ *and* $\mathscr{D}(\overline{V}) \leq \overline{V}$.

Assumption 8. *(i) Dom* V^* *is closed subset of* $\mathbb{S} \times \mathbb{R}^N$. *(ii)* $\underline{V}$ *and* $\overline{V}$ *are convex and lower semicontinuous on* $\mathbb{S} \times \mathbb{R}^N$. *(iii)* $\underline{V}$ *is level bounded. (iv) Dom* $\overline{V}$ *= Dom* V^* *and for each* $s \in \mathbb{S}$, $(s, 0) \in$ *Dom* $\overline{V}$.

Modify the definition of $\mathscr{V}$:

$$\mathscr{V} = \{V : \mathbb{S} \times \mathbb{R}^N \to \mathbb{R} \cup \{\infty\} | V \text{ is convex and lower semicontinuous and } \underline{V} \leq V \leq \overline{V}\}. \quad (43)$$

By Assumption 8 (iii) all functions in $\mathscr{V}$ are level bounded and by Assumption 8 (iv) all have Dom V = Dom V^* with for each $s \in \mathbb{S}$, $V(s,0) < \infty$. Define $\mathscr{T}(V)(s,\mu,\eta) = \sum_{\mathbb{S}}[\nu_{s,s'}(\mu,\eta) + \beta V(s',\mu'_{s,s'}(\mu,\eta))]Q(s,s')$ and assume the following.

Assumption 9. *Let $B : \mathbb{S} \times \mathbb{R}^N \times \mathbb{R}^M_+ \to \mathbb{R} \cup \{\infty\}$, $B(s,\mu,\eta) := \sum_{s'} \nu_{s,s'}(\mu,\eta)Q(s,s')$, be bounded below. If $V \in \mathscr{V}$, then $\mathscr{T}(V)$ is lower level bounded locally in μ.*

We have the following result.

Proposition 9. *Assume $V \in \mathscr{V}$ and that Assumptions 3, 4, 5′, 8 and 9 hold. Then $\mathscr{D}(V)$ is inf-proper and lower semicontinuous.*

Proof. By Assumption 9 and the definition of $\mathscr{V}$, both $B(s,\mu,\eta)$ and $V \in \mathscr{V}$ are bounded below. Thus, $\mathscr{T}(V)$ is bounded below. Since for each $s \in \mathbb{S}$, $B(s,0,0) = \sup_{y \in C(s)} f(s,y) < \infty$ and $V(s,0) < \infty$, we have for all $s \in \mathbb{S}$, $\mathscr{T}(V)(s,0) < \infty$. Hence, $\mathscr{T}(V)$ is inf-proper. By Aliprantis and Border (1999), p. 538, and Assumptions 3 and 4, ν is lower semicontinuous and so, since V is lower semicontinuous, $\mathscr{T}(V)$ is lower semicontinuous. Since V is level bounded, Assumption 9 implies that $\mathscr{T}(V)$ is level bounded locally in μ. Then, from Rockafellar and Wets (1998), Theorem 1.17, p. 16 and Theorem 3.31, p. 93, $\mathscr{D}(V)$ is inf-proper and lower semicontinuous. □

The following is an immediate implication of Proposition 9 and assumptions.

Lemma 2. *Let Assumptions 3, 4, 5′, 8 and 9 hold. If $V \in \mathscr{V}$, then $\mathscr{D}(\mathscr{V}) \in \mathscr{V}$.*

Proof. $\mathscr{D}$ is easily shown to be monotone. Thus, if $\underline{V} \leq V \leq \overline{V}$, then, using Assumption 5′, $\underline{V} \leq \mathscr{D}(\underline{V}) \leq \mathscr{D}(V) \leq \mathscr{D}(\overline{V}) \leq \overline{V}$. $\mathscr{D}$ preserves convexity and lower semicontinuity by the assumptions and Proposition 9. Hence, $\mathscr{D}(V) \in \mathscr{V}$. □

As before $\mathscr{V}$ is a complete metric space and Proposition 5 holds. Consequently, the following modified version of Theorem 1 obtains.

Theorem 1′. *Let Assumptions 3, 4, 5′, 8 and 9 hold and $\mathscr{V}$ be defined as in (43). V^* is the unique fixed point of $\mathscr{D}$ in $\mathscr{V}$. Also, there is an $\alpha \in [0,1)$ such that for any $V \in \mathscr{V}$, $\mathscr{D}^n(V) \xrightarrow{d} V^*$ with $d(\mathscr{D}^n(V), V^*) \leq \alpha^n d(V,V^*) \leq \alpha^n d(\overline{V},\underline{V})$.*

In many instances the main difficulty in applying Theorem 1′ is the verification of Assumption 9. For the hidden information problem with i.i.d. shocks and utility unbounded below, Dom $V^* = \mathbb{R}_+$ (and not $\mathbb{R}$). Defining $\underline{V}$ and $\overline{V}$ equal to the value functions from no insurance and full insurance problems and $\underline{\underline{V}}$ as before, it is easy to verify Assumptions 5′ and 8. Letting $\overline{y} > 0$ (after possible renormalization of the agent's utility function), the first part of Assumption 9 is assured, since then $0 \in Y$ and $B(s,\mu,\eta) \geq f(s,0) > -\infty$. The second part of Assumption 9 is verified by checking that if $V \in \mathscr{V}$ and, hence, is bounded below and level bounded, at any (s,μ), allowing $\|\eta\|$ to become arbitrarily large causes $B(s,\mu,\eta)$ or $V(s',\mu'_{s,s'}(\mu,\eta))$ to become arbitrarily large. Thus, Theorem 1′ is applicable.

References

Aliprantis, C. and K. Border (1999). *Infinite Dimensional Analysis*. Springer-Verlag.

Atkeson, A. and R. Lucas (1992). Efficient distributions with private information. *Review of Economic Studies 59*, 427–453.

Boyd, J. (1990). Recursive utility and the ramsey problem. *Journal of Economic Theory 50*, 326–345.

Cole, H. and F. Kubler (2010). Recursive contracts, lotteries and weakly concave pareto sets. University of Pennsylvania.

da Rocha, V. M. and Y. Vailakis (2010). Existence and uniqueness of a fixed point for local contractions. *Econometrica 78*, 1127–1141.

Duran, J. (2000). Dynamic programming with unbounded returns. *Economic Theory 15*, 332–350.

Farhi, E. and I. Werning (2007). Inequality and social discounting. *Journal of Political Economy 115*, 365–402.

Hernandez-Lerma, O. and J. Laserre (1999). *Further Topics on Discrete-Time Markov Control Processes*. Springer-Verlag.

Luenberger, D. (1969). *Optimization by Vector Space Methods*. New York, John Wiley & Sons.

Marcet, A. and R. Marimon (1999). Recursive contracts. Working paper.

Marcet, A. and R. Marimon (2011). Recursive contracts. Working Paper.

Matkowski, J. and A. Nowak (2011). On discounted dynamic programming with unbounded returns. *Economic Theory 46*, 455–474.

Mele, A. (2010). Repeated moral hazard and recursive lagrangeans. Working paper.

Messner, M. and N. Pavoni (2004). On the recursive saddle point method. IGIER Working Paper 255.

Messner, M., N. Pavoni, and C. Sleet (2011a). Dual recursive methods. CMU Working Paper.

Messner, M., N. Pavoni, and C. Sleet (2011b). Recursive methods for incentive problems. CMU Working Paper.

Rincon-Zapatero, J. and C. Rodriguez-Palermo (2003). Existence and uniqueness of solutions to bellman equation in the unbounded case. *Econometrica 71*, 1519–1555.

Rincon-Zapatero, J. and C. Rodriguez-Palermo (2009). Correction to existence and uniqueness of solutions to bellman equation in the unbounded case. *Econometrica 77*, 317–319.

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

Rockafellar, T. and R. Wets (1998). *Variational Analysis*. Springer-Verlag.

Sleet, C. and S. Yeltekin (2009). Social credibility, social patience and long run inequality. Carnegie Mellon University.

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

Wessels, J. (1977). Markov programming by successive approximation with respect to weighted supremum norms. *Journal of Mathematical Analysis and Applications 58*, 326–355.