Econometrica Supplementary Material

ONLINE APPENDIX
THE DUAL APPROACH TO RECURSIVE OPTIMIZATION: THEORY AND EXAMPLES

(*Econometrica*, Vol. 86, No. 1, January 2018, 133–172)

NICOLA PAVONI
Department of Economics, Bocconi University, IGIER, and IFS

CHRISTOPHER SLEET
Tepper School of Business, Carnegie Mellon University

MATTHIAS MESSNER
Center for Macroeconomic Research, Universität zu Köln

APPENDIX A: NUMERICAL IMPLEMENTATION

THIS SECTION DESCRIBES how to implement the recursive dual approach numerically. Under the conditions of Theorem 2, the dual Bellman operator is a contraction and, consequently, it is natural to calculate D^* via value iteration. Numerical approximation of candidate dual value functions is facilitated by their sub-linearity and the simplicity of their domain. The dual Bellman involves an (outer) minimization over a set of multipliers; these multipliers are passed to (and "coordinate") a family of simple (inner) maximizations over current actions and states. Additive separability in the objective may be exploited to decompose the inner maximizations into a family of simpler maximizations that in parametric settings often have analytical solutions.

Dual Value Function Approximation. Numerical implementation of a value function iteration algorithm requires approximations to candidate value functions. Our implementation exploits the sub-linearity of dual value functions and uses a piecewise linear approximation (on the spherical domain $\mathcal{C}$). Piecewise linear approximations to value functions defined on spheres were first applied in economics by Judd, Yeltekin, and Conklin (2003). We apply their approximation procedure to our setting.[31] Recall that under the conditions of Theorem 2, the domain for the dual Bellman operator may be identified with an interval of functions $D : \mathcal{S} \times \mathcal{Y} \to \mathbb{R}$, each of which is sub-linear in its second argument. As noted, these functions are fully determined on $\mathcal{S} \times \mathcal{C}$ (or a subset thereof). Moreover, their sub-linearity implies that, for all $y \in \mathcal{C}$,[32]

$$D(s, y) = \max_{r \in \mathcal{V}} \bigl\{ r \cdot y \mid \forall y' \in \mathcal{C}, r \cdot y' \leq D\bigl(s, y'\bigr) \bigr\}. \tag{A.1}$$

Nicola Pavoni: pavoni.nicola@gmail.com
Christopher Sleet: csleet@andrew.cmu.edu
Matthias Messner: matthias.messner@unibocconi.it

[31] Judd, Yeltekin, and Conklin (2003) used this approach to approximate the support function of a payoff set in a repeated game; we use it to approximate the recursive dual value function. In other aspects, our (recursive dual) formulation is different from that of Judd, Yeltekin, and Conklin (2003). Alternative approaches to approximation on spherical domains are described in Sloan and Womersley (2000).

[32] For this and other properties of sub-linear functions used below, see Florenzano and Van (2001).

 https://doi.org/10.3982/ECTA11905

Given such a function D and a set of N distinct points $\hat{\mathcal{C}}^N := \{y_n\}_{n=1}^N \subset \mathcal{C}$, define the approximation $\hat{D}^N$ as, for each $(s, y) \in \mathcal{S} \times \mathcal{C}$,

$$\hat{D}^N(s, y) := \max_{r \in \mathcal{V}} \{ r \cdot y | \forall y_n \in \hat{\mathcal{C}}^N, r \cdot y_n \leq D(s, y_n) \}. \tag{A.2}$$

Since the constraint set in problem (A.2) is less restrictive than that in (A.1),

$$D \leq \hat{D}^N,$$

with equality at each $(s, y_n) \in \mathcal{S} \times \hat{\mathcal{C}}^N$. In addition, the approximation $\hat{D}^N$ remains sublinear, is summarized by $\{s, y_n, \hat{D}^N(s, y_n)\}_{n=1}^N$, and is easily evaluated by solving the simple linear programming problem in (A.2). Let $\{\hat{\mathcal{C}}^N\}$, $N \geq 1$, be a sequence of subsets of $\mathcal{C}$ such that (i) for all N, $\hat{\mathcal{C}}^N \subset \hat{\mathcal{C}}^{N+1}$ and (ii) $\hat{\mathcal{C}}^\infty = \bigcup_N \hat{\mathcal{C}}^N$ is dense in $\mathcal{C}$.[33] It is readily verified that the corresponding sequence of approximating functions $\hat{D}^N(s, \cdot)$ converges pointwise to $D(s, \cdot)$ from above.[34] Moreover, by Dini's theorem,[35] it converges uniformly on $\mathcal{C}$ and, hence, in the Thompson-like metric d to $D(s, \cdot)$. In practical applications, we use (hyper)spherical coordinates to represent points in $\hat{\mathcal{C}}^N$. The corresponding Cartesian coordinates of points in this grid are recovered from spherical coordinates $\{\phi_n^r\}$ according to the formulas $y_n^1 = \cos(\phi_n^1)$, for $j = 2, \ldots, n_K + n_V - 1$, $y_n^j = \cos(\phi_n^j) \prod_{r=1}^{j-1} \sin(\phi_n^r)$, and $y_n^{n_K+n_V} = \prod_{r=1}^{n_K+n_V-1} \sin(\phi_n^r)$.

The approximation procedure described above may be integrated into the dual value iteration to give the $\nu + 1$-iteration step[36]

$$\begin{aligned} &\forall (s, y_n) \in \mathcal{S} \times \hat{\mathcal{C}}^N, \\ &\hat{\mathcal{B}}(\hat{D}_\nu^N)(s, y_n) = \inf_{q \in \mathcal{Q}} \sup_{p \in \mathcal{P}} J(s, y_n; q, p) + \delta \sum_{s' \in \mathcal{S}} D_\nu^N(s', y'(s')) \pi(s'|s). \end{aligned} \tag{A.3}$$

Optimization. The inner supremum operation in (A.3) results in the indirect current dual function $J^*(s, y_n; q) = \sup_{p \in \mathcal{P}} J(s, y_n; q, p)$. Additive separability of the function J across different components of p can often be exploited to break the supremum down into separate optimizations over the components of p which can be run in parallel or in some cases solved analytically. In these latter cases, no explicit numerical maximization over primal choices is needed. Once the inner suprema are solved, an indirect objective

[33]For example, the set of points in $\mathcal{C}$ with rational coordinates is dense in $\mathcal{C}$; see Schmutz (2008).

[34] It clearly converges at all points in $\hat{\mathcal{C}}^\infty$. Choose a point $y \in \mathcal{C}$. Let $\{y_r^1\}$ and $\{y_r^2\}$ be two sequences in $\bigcup_N \hat{\mathcal{C}}^N$ converging to y and such that $y = \lambda_r a_r y_r^1 + (1 - \lambda_r) b_r y_r^2$, with $\lambda_r \in (0, 1)$, $a_r, b_r \in \mathbb{R}_+$, and $a_r, b_r \downarrow 1$, that is, $a_r y_r^1$ and $b_r y_r^2$ lie either side of y on the tangent to $\mathcal{C}$ passing through y. There is a sequence $\{N_r\}$ such that $\hat{D}^{N_r}(s, y_r^1) = D(s, y_r^1)$ and $\hat{D}^{N_r}(s, y_r^2) = D(s, y_r^2)$. By the sub-linearity of $D(s, \cdot)$ and each $\hat{D}^{N_r}(s, \cdot)$, we have $D(s, y) \leq D^{N_r}(s, y) \leq \lambda_r \hat{D}^{N_r}(s, a_r y_r^1) + (1 - \lambda_r) \hat{D}^{N_r}(s, b_r y_r^2) = \lambda_r a_r \hat{D}^{N_r}(s, y_r^1) + (1 - \lambda_r) b_r \hat{D}^{N_r}(s, y_r^2) = \lambda_r a_r D(s, y_r^1) + (1 - \lambda_r) b_r D(s, y_r^2)$. Since $D(s, \cdot)$ is real-valued and convex, it is continuous at all interior points; by linear homogeneity, $D(s, \cdot)$ is continuous throughout $\mathcal{Y}$, hence $y_r^1 \to y$, $y_r^2 \to y$, and $a_r, b_r \downarrow 1$, it follows that the last term in the string of inequalities converges to $D(s, y)$. Thus, the sequence of functions converges pointwise on $\mathcal{C}$ and by the positive homogeneity of the functions on $\mathcal{Y}$ as well.

[35]See Chapter 2, Aliprantis and Border (2006) for a statement and proof of Dini's theorem.

[36]With some simplification if the problem is quasilinear.

over multipliers is obtained and (A.3) becomes

$$\forall (s, y_n) \in \mathcal{S} \times \hat{\mathcal{C}}^N, \quad \hat{\mathcal{B}}\big(\hat{D}_\nu^N\big)(s, y_n) = \inf_{q \in \mathcal{Q}} J^*(s, y_n; q) + \delta \sum_{s' \in \mathcal{S}} D_\nu^N\big(s', y'(s')\big)\pi\big(s'|s\big). \quad \text{(A.4)}$$

The objective in (A.4) is convex (even if the underlying problem is not), but it is not smooth.[37] There are many optimization procedures for non-smooth, convex dual problems (e.g., sub-gradient algorithms, cutting plane algorithms, and so forth[38]). These may be used to solve the problems (A.4). An alternative approach developed by Necoara and Suykens (2008) is to smooth the dual problem through the addition of strongly concave (prox) functions to the objective in (A.2) (and, if necessary, the objective J in the inner sup problems). In our calculations, we follow Necoara and Suykens (2008) by adding terms $c_v\|r\|^2$ to the objective in (A.2) and allowing $c_v \to 0$ with successive iterations. We use the optimizer SNOPT to solve these (smoothed) optimizations.

APPENDIX B: THE LIMITED COMMITMENT EXAMPLE

This section collects details of and extensions to the results of Section 2.

B.1. *Numerical Method*

The numerical approach is outlined in Section A. We apply this approach to the transformed version of the problem described in Section 2. Inner maximizations are solved analytically when possible and the indirect payoffs $J^*(s, y; q)$ are substituted directly into (5). On the νth application of the Bellman operator, the following family of minimizations is solved, for $s \in \mathcal{S}$ and $y_n \in \hat{\mathcal{C}}_+^N$:

$$D_{\nu+1}^N(s, y_n) = \inf_q J^*(s, y_n; q) + \delta \sum_{s' \in \mathcal{S}} D_\nu^N\big(s', y'(s')\big)\pi\big(s'|s\big), \quad \text{(B.1)}$$

where J^* is defined as in (7) and $\hat{\mathcal{C}}_+^N$ is a finite subset of $\mathcal{C}_+ = \mathcal{C} \cap \mathbb{R}_+^{n_I}$ and is represented as a grid of points in spherical coordinates (either $\{\phi_n\} \subset [0, \pi)$ if $n_I = 2$ or $\{\phi_n^1, \phi_n^2\} \in [0, \pi)^2$ if $n_I = 3$). In the $n_I = 2$ case, the spherical coordinate gives the (Pareto) weights on agents 1 and 2 according to $y^1 = \cos\phi$ and $y^2 = \sin\phi$; in the $n_I = 3$ case, ϕ^1 gives the weight on agent 1 relative to agents 2 and 3, while ϕ^2 gives the weight on agent 2 relative to agent 3.

B.2. *Construction of Bounding Functions*

Application of Theorem 2 (and the proof that $\mathcal{B}$ is a contraction) requires the definition of bounding value functions $\underline{\underline{D}}$, $\underline{D}$, and $\overline{D}$. In this subsection, bounding functions satisfying Assumption 3 are obtained for the transformed limited commitment problem. Let $\overline{v} := \max_{\mathcal{S}} \frac{\gamma(s)^{1-\sigma}}{1-\sigma}$ and $\underline{v} = \frac{(\underline{a}/2)^{1-\sigma}}{1-\sigma}$, where $\underline{a}$ is the nonnegative (positive if $\sigma > 1$) lower bound on agent consumptions. Assume an $\tilde{a} \in \mathcal{A}^{n_S}$ and let $\xi > 0$ be such that for each $s \in \mathcal{S}$,

[37] J^* may be smooth if it is obtained from component problems with strictly concave objectives and concave constraint functions. However, our approximation procedure implies that $\hat{D}^N$ is non-smooth.

[38] Good references for such methods include Bertsekas (2003) and Ruszczyński (2006).

$\gamma(s) > \sum_{i\in\mathcal{I}} \tilde{a}^i(s)$, and for each $s \in \mathcal{S}$ and $i \in \mathcal{I}$,

$$\begin{aligned}\overline{v} - \xi &\geq \frac{1-\delta}{1-\sigma}\big[\tilde{a}^i(s)\big]^{1-\sigma} + \delta\overline{v} \\ &> \frac{1-\delta}{1-\sigma}\big[\tilde{a}^i(s)\big]^{1-\sigma} + \frac{\delta}{1-\sigma}\bigg\{\sum_{s'\in\mathcal{S}}\big[(1-\sigma)w^i(s')\big]^{\theta}\pi(s'|s)\bigg\}^{\frac{1}{\theta}} > w^i(s) + \xi,\end{aligned} \tag{B.2}$$

where $\theta := \frac{1-\rho}{1-\sigma}$. Set

$$\overline{D}(s, y) = \sum_{i\in\mathcal{I}} y^i\varphi^i(y^i, s), \quad \varphi^i(y^i, s) = \begin{cases}\overline{v} & \text{if } y^i \geq 0,\\ \underline{v} & \text{else if } y^i < 0,\end{cases}$$

and

$$\underline{\underline{D}}(s, y) = \sum_{i\in\mathcal{I}}\big\{y^i\psi^i(y^i, s) + |y^i|\xi\big\}, \quad \psi^i(y^i, s) = \begin{cases} w^i(s) & \text{if } y^i \geq 0,\\ \overline{v} & \text{else if } y^i < 0.\end{cases}$$

It is immediate that $\overline{D}$ is continuous and that $\underline{\underline{D}}$ is continuous and positively homogeneous. It is also easy to see that $\overline{D} \geq D^*$: while the supremum operations defining D^* are restricted by feasibility and default constraints, $\overline{D}$ gives the maximal weighted payoff subject only to the restriction that payoffs remain within $\mathcal{V}$. In addition, it follows from (B.2) that $\underline{\underline{D}} \leq D^*$.

We verify that for $\varepsilon > 0$, $\mathcal{B}(\overline{D}) \leq \overline{D}$ and $\mathcal{B}(\underline{\underline{D}}) > \underline{\underline{D}} + \varepsilon$ on $\mathcal{S} \times \mathcal{C}$. $\mathcal{B}(D)$ is given by, for all $(s, y) \in \mathcal{S} \times \mathcal{Y}$,

$$\begin{aligned}&\mathcal{B}(D)(s, y) \\ &\quad= \inf_{\mathcal{Q}}\sup_{\mathcal{P}}\sum_{i\in\mathcal{I}}(y^i + m^i)\bigg\{\frac{1-\delta}{1-\sigma}\big[a^i\big]^{1-\sigma} + \frac{\delta}{1-\sigma}\bigg\{\sum_{s'\in\mathcal{S}}\big[(1-\sigma)v'^{,i}(s')\big]^{\theta}\pi(s'|s)\bigg\}^{\frac{1}{\theta}}\bigg\} \\ &\qquad - \sum_{i\in\mathcal{I}} m^i w^i(s) - m^{n_I+1}\bigg(\sum_{i\in\mathcal{I}} a^i - \gamma(s)\bigg) - \delta\sum_{s'\in\mathcal{S}}\sum_{i\in\mathcal{I}} y'^{,i}(s')v'^{,i}(s')\pi(s'|s) \\ &\qquad + \delta\sum_{s'\in\mathcal{S}} D\big(s', y'(s')\big)\pi(s'|s).\end{aligned} \tag{B.3}$$

Setting $D = \overline{D}$, using the definition of $\underline{v}$ and $\overline{v}$, and noting that the dual variables (m, y') can always be chosen equal to 0 in the infimum and $\overline{D}(s, 0) = 0$, we have $\mathcal{B}(\overline{D})(s, y) \leq \overline{D}(s, y)$. Finally, we show $\mathcal{B}(\underline{\underline{D}}) > \underline{\underline{D}} + \varepsilon$ on $\mathcal{S} \times \mathcal{C}$. Given $y' = \{y'^{,i}(s')\}$, define $\psi(y') = \{\psi^i(y'^{,i}(s'), s')\}_{(i,s')\in\mathcal{I}\times\mathcal{S}}$. Setting $D = \underline{\underline{D}}$ and noting that for any s and choice of (m, y'), the pair $(\tilde{a}(s), \psi(y'))$ is a feasible choice for the supremum with respect to both the resource and no default constraints, we have

$$\begin{aligned}&\mathcal{B}(\underline{\underline{D}})(s, y) \\ &\quad\geq \inf_{q\in\mathcal{Q}}\sum_{i\in\mathcal{I}}(y^i + m^i)\bigg\{\frac{1-\delta}{1-\sigma}\big[\tilde{a}^i(s)\big]^{1-\sigma} + \frac{\delta}{1-\sigma}\bigg\{\sum_{s'\in\mathcal{S}}\big[(1-\sigma)\psi^i\big(y'^{,i}(s'), s'\big)\big]^{\theta}\pi(s'|s)\bigg\}^{\frac{1}{\theta}}\bigg\}\end{aligned}$$

$$- \sum_{i\in\mathcal{I}} m^i w^i(s) - m^{n_I+1}\left(\sum_{i\in\mathcal{I}} \tilde{a}^i(s) - \gamma(s)\right)$$

$$\geq \inf_{y'} \sum_{i\in\mathcal{I}} y^i \left\{ \frac{1-\delta}{1-\sigma}\left[\tilde{a}^i(s)\right]^{1-\sigma} + \frac{\delta}{1-\sigma}\left\{\sum_{s'\in\mathcal{S}}\left[(1-\sigma)\psi^i\left(y^{\prime,i}(s'), s'\right)\right]^{\theta}\pi\left(s'|s\right)\right\}^{\frac{1}{\theta}}\right\},$$

where the first inequality follows from the replacement of the sup with the choices $(\tilde{a}(s), \psi(y'))$. The second inequality uses the feasibility of these choices and thus the fact that $m = 0$ is minimizing. Now, using the additive separability across agents, each agent i can be analyzed separately. If $y^i \geq 0$, then

$$\inf_{y^{\prime,i}(\cdot)} y^i \left\{ \frac{1-\delta}{1-\sigma}\left[\tilde{a}^i(s)\right]^{1-\sigma} + \frac{\delta}{1-\sigma}\left\{\sum_{s'\in\mathcal{S}}\left[(1-\sigma)\psi^i\left(y^{\prime,i}(s'), s'\right)\right]^{\theta}\pi\left(s'|s\right)\right\}^{\frac{1}{\theta}}\right\}$$

$$\geq y^i \left\{ \frac{1-\delta}{1-\sigma}\left[\tilde{a}^i(s)\right]^{1-\sigma} + \frac{\delta}{1-\sigma}\left\{\sum_{s'\in\mathcal{S}}\left[(1-\sigma)w^i\left(s'\right)\right]^{\theta}\pi\left(s'|s\right)\right\}^{\frac{1}{\theta}}\right\}$$

$$\geq y^i\left(w^i(s) + \xi\right),$$

with the inequality strict if $y^i > 0$. Note that if $\sigma > 1$, then $0 \leq (1-\sigma)\psi^i(y^{\prime,i}(s'), s') \leq (1-\sigma)w^i(s')$, $\forall i, s$, and hence $\{\sum_{\mathcal{S}}[(1-\sigma)\psi^i(y^{\prime,i}(s'), s')]^{\theta}\pi(s'|s)\}^{\frac{1}{\theta}} \leq \{\sum_{\mathcal{S}}[(1-\sigma)w^i(s')]^{\theta} \times \pi(s'|s)\}^{\frac{1}{\theta}}$, which implies the above inequality when both sides are multiplied by the negative number $\frac{\delta}{1-\sigma}$. The last inequality holds by our assumption on $\tilde{a}^i(s)$. Similarly, if $y^i < 0$, then

$$\inf_{y^{\prime,i}} y^i \left\{ \frac{1-\delta}{1-\sigma}\left[\tilde{a}^i(s)\right]^{1-\sigma} + \frac{\delta}{1-\sigma}\left\{\sum_{s'\in\mathcal{S}}\left[(1-\sigma)\psi^i\left(y^{\prime,i}(s'), s'\right)\right]^{\theta}\pi\left(s'|s\right)\right\}^{\frac{1}{\theta}}\right\}$$

$$\geq y^i \left\{ \frac{1-\delta}{1-\sigma}\left[\tilde{a}^i(s)\right]^{1-\sigma} + \delta\overline{v}\right\} > y^i(\overline{v} - \xi).$$

Consider now the auxiliary function

$$\tilde{D}(s, y) = \sum_{i\in\mathcal{I}}\left\{y^i\tilde{\psi}^i\left(y^i, s\right) + \left|y^i\right|\xi\right\},$$

where

$$\tilde{\psi}^i\left(y^i, s\right) = \begin{cases} \frac{1-\delta}{1-\sigma}\left[\tilde{a}^i(s)\right]^{1-\sigma} + \frac{\delta}{1-\sigma}\left\{\sum_{s'\in\mathcal{S}}\left[(1-\sigma)w^i\left(s'\right)\right]^{\theta}\pi\left(s'|s\right)\right\}^{\frac{1}{\theta}} & \text{if } y^i \geq 0, \\ \frac{1-\delta}{1-\sigma}\left[\tilde{a}^i(s)\right]^{1-\sigma} + \delta\overline{v} & \text{else if } y^i < 0. \end{cases}$$

$\tilde{D}(s, y)$ is clearly a continuous function. Our preceding derivations show that for all $(s, y) \in \mathcal{S}\times\mathcal{C}$, $\mathcal{B}(\underline{\underline{D}})(s, y) \geq \tilde{D}(s, y) > \underline{\underline{D}}(s, y)$. The continuity of each $\underline{\underline{D}}(s, \cdot)$ and $\tilde{D}(s, \cdot)$ and the compactness of $\mathcal{C}$ then imply that there is an $\varepsilon > 0$ such that, for all $(s, y) \in \mathcal{S}\times\mathcal{C}$, $\mathcal{B}(\underline{\underline{D}})(s, y) \geq \tilde{D}(s, y) > \underline{\underline{D}}(s, y) + \varepsilon$, and hence $\mathcal{B}(\underline{\underline{D}})(s, y) \geq \underline{\underline{D}}(s, y) + \varepsilon$ as required.

B.3. *Numerical Calculations*

This subsection reports additional calculations for the two and three agent cases.

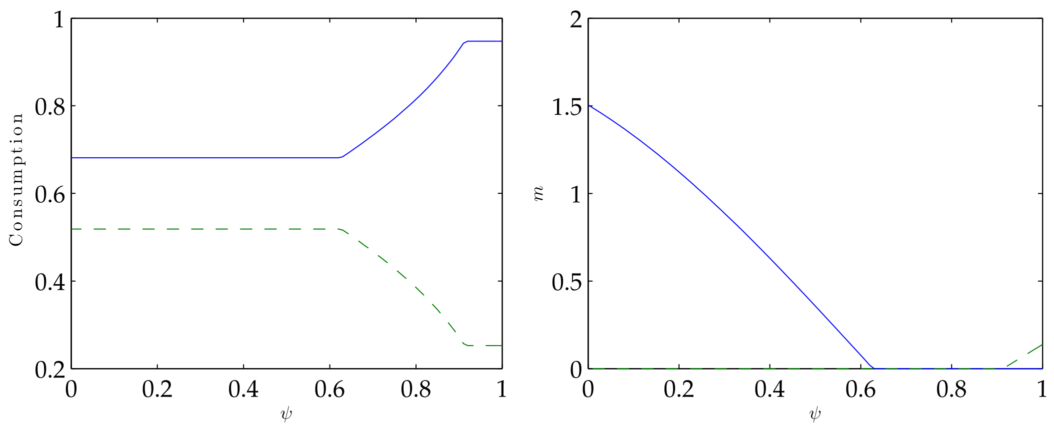
B.3.1. *Two Agents*

In the main text, results for the case $\sigma = 1.5$ and $\rho = 5$ are reported. Values of ρ that are smaller and closer to σ result in more volatility in consumption and a muting of the dynamics that occur when none of the no default constraints are binding. In the limiting case of expected utility preferences $\rho = \sigma$, these dynamics disappear completely. Figures 1 and 2 illustrate for the case $\sigma = \rho = 1.5$.

B.3.2. *Three Agents*

We now show computed results from a three agent economy. The preference parameters are set to $\sigma = 1.5$, $\rho = 5$, and $\delta = 0.8$. In shock state s, agent s has an outside option equal to the utility from a steady endowment stream of 40% of the total endowment; agents $s' \neq s$ have outside options equal to the utility from a steady endowment stream of 10% of the total endowment. These values preclude full risk sharing. It is convenient to plot policies as functions of spherical coordinates (ϕ_1, ϕ_2). The corresponding costates or "Pareto weights" on agents are $y^1 = \cos\phi_1$, $y^2 = \sin\phi_1 \cos\phi_2$, and $y^3 = \sin\phi_1 \sin\phi_2$. Thus, higher values of ϕ_1 imply less weight on agent 1's utility and more on agent 2's and agent 3's, while higher values of ϕ_2 imply less weight on agent 2's utility and more on agent 3's (with no change in the weight on agent 1's). Figure 3 shows the computed optimal dual value function and the Thompson metric distance between successive iterates, illustrating the geometric convergence of the value iteration.

Figure 4 shows calculated policy functions. Panel (a) of the figure displays the optimal multiplier m^2 for agent 2 in shock state 2 (in which agent 2 has a high outside option and agents 1 and 3 low ones). The weight y^2 is small and agent 2's incentive multiplier correspondingly large when the spherical coordinates ϕ_1 and ϕ_2 are, respectively, small



(a) Consumption (b) Multipliers

FIGURE 1.—Solid lines give agent 1's policy; dashed lines, agent 2's policy. Policies are given as functions of agent 1's normalized costate and for $s = 1$. Panel (a) shows agent consumption; panel (b), multipliers on the commitment constraints.

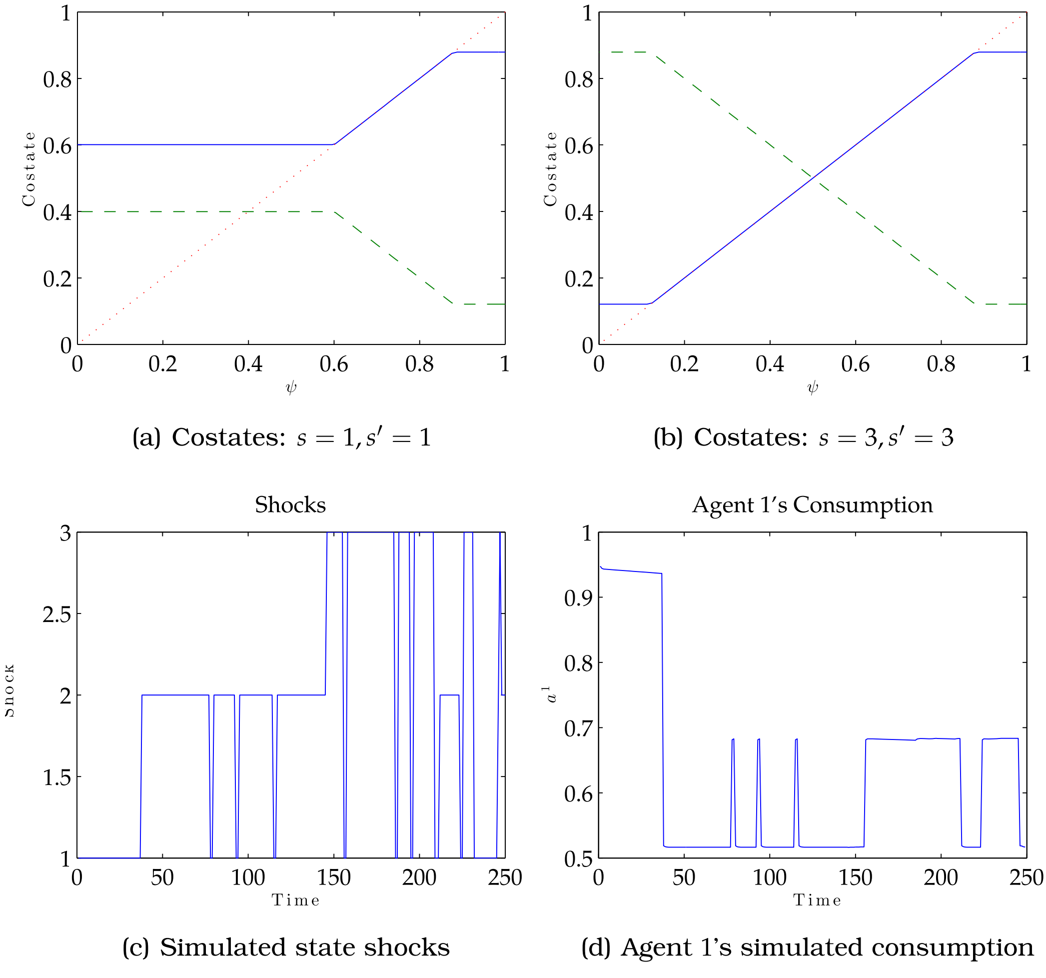
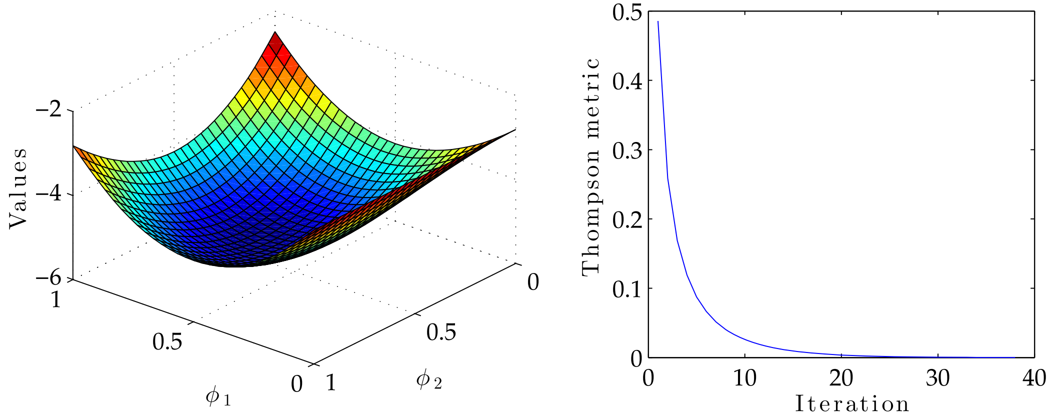


FIGURE 2.—Panels (a) and (b) show the (normalized) costates associated with remaining in state 1 and state 3, respectively, as a function of the initial value of the costate ψ. Solid lines give agent 1's policy; dashed lines, agent 2's policy. The 45 degree line is illustrated with dots. Policies are given as functions of agent 1's normalized costate and for $s = 1$. Panels (c) and (d) illustrate a 250 period simulation of agent 1's consumption which displays the usual 'memoryless property on the ergodic set'. When a change of state s leads a different agent's no default constraint to bind, an adjustment in agent 1's consumption occurs. Thus, this agent's consumption bounces between two (history independent) levels; it remains constant whenever the economy remains in the same state or transitions into $s = 3$.

and large. Then, the combination of a low costate and a high outside option imply that additional consumption must be given to agent 2 now and in the future to keep her inside the risk sharing arrangement. Panels (b), (c), and (d) of Figure 4 show the consumption of agents 1, 2, and 3 as functions of the costates again given $s = 2$. Each agent's consumption rises in areas of the state space corresponding to a higher (Pareto) weighting. Agent 2's consumption is sustained above 0.5 by her binding incentive constraint, while the consumption of agents 1 and 3 decreases towards 0.2 as their (Pareto) weights decrease to zero.



(a) Value function (b) Convergence

FIGURE 3.—Panel (a) shows the planner's dual value function in shock state $s = 1$ at the terminal iteration. Panel (b) shows the Thompson metric distance between iterates.

Panels (a) and (b) of Figure 5 show "quiver plots" indicating the direction in which the spherical coordinates describing costates are updated if the economy remains in shock state $s = 1$ (panel a) and $s = 2$ (panel b). Consider panel (a). Recall that high values of ϕ_1 imply that the costate on agent 1 is low. Agent 1's no default constraint is then binding, her multiplier is positive, and her costate is raised. The spherical coordinate ϕ_1 is correspondingly reduced (placing less weight on agents 2 and 3 and more on agent 1). This is indicated by a left pointing arrow at high ϕ_1 values (on the right-hand side of the plot). Low values for ϕ_1 and high values for ϕ_2 imply that the costate on agent 2 is low. If it is low enough, then even in state 1 (when agent 2's outside option is low), agent 2's outside option binds. Hence, agent 2's costate is increased. Spherical coordinate ϕ_1 is correspondingly increased and ϕ_2 decreased and the arrows in the top left-hand corner of the plot point down and inwards. Similar reasoning holds with respect to the bottom left-hand corner of the plot, where agent 3's costate is low and outside option binds and the arrows point up and inwards. In the dotted region in the center left of the plot, no incentive constraints bind. Here, very small adjustments to costates occur that stem from the early resolution of uncertainty structure of preferences and the force for equality that it imparts. Panel (b) shows the adjustments in spherical coordinates when the economy remains in shock state $s = 2$ and agent 2's outside option is large. It has an analogous interpretation to panel (a).

Panels (c) and (d) show a simulation of shocks and agent 1's consumption. Consumption remains relatively stable if the economy persists in a given shock state. But large adjustments occur when a transition into or out of shock $s = 1$ occurs (and agent 1's outside option abruptly changes relative to the other agents' outside options).

APPENDIX C: PROOF OF PROPOSITION 8

We first establish existence of a saddle point with summable multipliers for an abstract problem with inequality constraints. We then relate this problem to a modified version of (P) (called (MP)). We associate a Lagrangian with (MP) and show that each primal

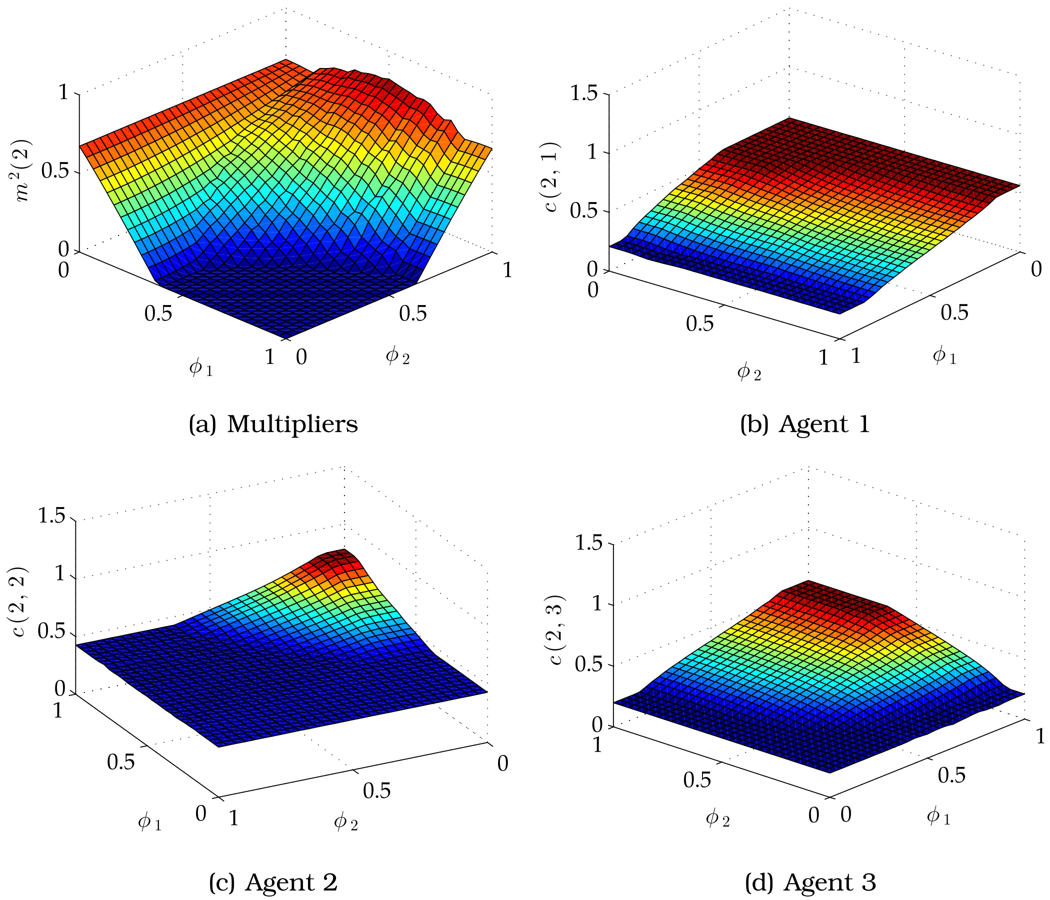


FIGURE 4.—Figures are drawn on a domain of costates expressed in spherical coordinates. Higher values of ϕ_1 place more weight on agents 2 and 3; higher values of ϕ_2 place more weight on agent 3. Panel (a) displays the incentive multiplier for agent 2 in shock $s = 2$. The remaining three panels display the consumption of agents in shock $s = 2$.

plan solving (MP) is part of a saddle point with a minimizing summable multiplier. Finally, we show that each solution to (P) defines a solution to (MP) and use the minimizing multiplier from (MP) to construct a minimizing multiplier and, hence, saddle point for (P).

The Abstract Problem. Consider

$$\sup f(x) \quad \text{s.t. } g(x) \geq 0, \tag{AP}$$

where $f : \ell^\infty \to \mathbb{R}$ and $g : \ell^\infty \to \ell^\infty$, with $g(x) = \{g_r(x)\}_{r=1}^\infty$ and each $g_r : \ell^\infty \to \mathbb{R}$. Associate the Lagrangian $\mathcal{L} : \ell^\infty \times \ell_+^{\infty,\star} \to \mathbb{R}$ with (AP), where[39]

$$\mathcal{L}(x, \lambda) := f(x) + \langle \lambda, g(x) \rangle.$$

[39]Define $\lambda \in \ell_+^{\infty,\star}$ if $\lambda \in \ell^{\infty,\star}$ and $\lambda \geq 0$, where $\lambda \geq 0 \iff \langle \lambda, y \rangle \geq 0\ \forall y \in \ell^\infty, y \geq 0$.

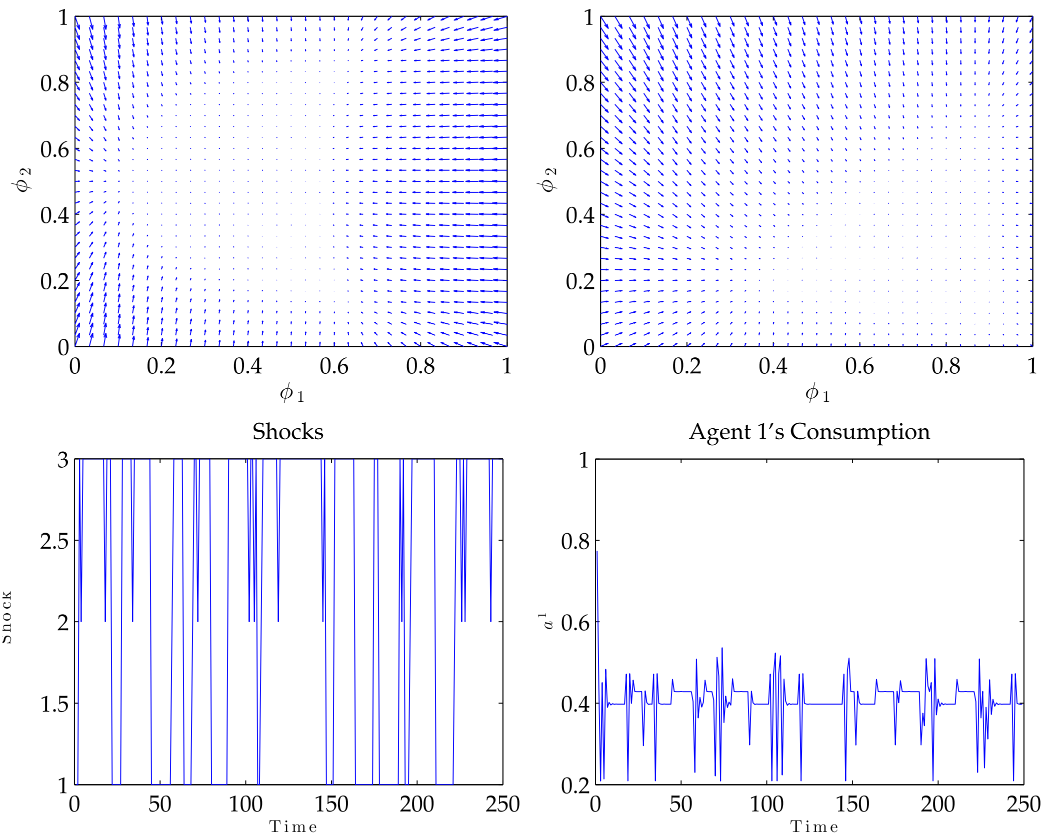


FIGURE 5.—Policy functions and simulations. Panels (a) and (b) show a quiver plot for the costate policy functions associated with remaining in states $s = 1$ and $s = 2$, respectively. The arrows in the plots indicate the direction in which the spherical coordinates describing costates are updated. Panels (c) and (d) show simulations of shocks and of the consumption of agent 1.

Given Assumption 7 below and the existence of a solution x^* to (AP), Theorem C.1 establishes the existence of a saddle point.

ASSUMPTION 7: (C) Concavity: f and g are concave. (S) Slater Condition: There is an $\hat{x}$ such that $\inf_r g_r(\hat{x}) > 0$.

THEOREM C.1—Saddle Point Existence for the Abstract Problem:

(i) *If x^* is feasible for* (AP) *and solves* $\max_{x\in\ell^\infty} \mathcal{L}(x, \lambda^*)$ *with* $\lambda^* \in \ell^{\infty,\star}$ *such that* $\lambda^* \geq 0$ *and* $\langle\lambda^*, g(x^*)\rangle = 0$, *then* x^* *solves* (AP).

(ii) *If Assumption* 7 *holds and* x^* *solves* (AP), *then there is a* $\lambda^* \in \ell^{\infty,\star}$ *such that* $\lambda^* \geq 0$ *and* $\langle\lambda^*, g(x^*)\rangle = 0$. *Moreover,* x^* *solves* $\max_{x\in\ell^\infty} \mathcal{L}(x, \lambda^*)$.

(iii) *If* (a) $\lambda^* \geq 0$ *and* (b) $\langle\lambda^*, g(x^*)\rangle = 0$, *then* x^* *is feasible for* (AP) *if and only if*

$$\mathcal{L}(x^*, \lambda^*) \leq \mathcal{L}(x^*, \lambda) \quad \forall\lambda \in \ell^{\infty,\star} \text{ with } \lambda \geq 0.$$

(iv) *If Assumption* 7 *holds and* x^* *solves* (AP), *then* $\mathcal{L}$ *has a saddle point in* $\ell^\infty \times \ell_+^{\infty,\star}$.

PROOF OF THEOREM C.1: (i) If x is feasible for (AP), then all nonlinear constraints must hold with inequality and $\langle \lambda^*, g(x)\rangle \geq 0$. Thus, for any feasible choice x: $f(x^*) = f(x^*) + \langle \lambda^*, g(x^*)\rangle = \mathcal{L}(x^*, \lambda^*) \geq \mathcal{L}(x, \lambda^*) = f(x) + \langle \lambda^*, g(x)\rangle \geq f(x)$. (ii) This proof is standard. To save space, we do not report it here, but refer the interested reader to Luenberger (1969), Theorem 1, pages 217–218. (iii) ($\Rightarrow$) If x^* is feasible, then $g(x^*) \geq 0$. Hence, for all $\lambda \geq 0$, $\langle \lambda, g(x^*)\rangle \geq 0$ and 0 is the infimum of $\langle \lambda, g(x^*)\rangle$ over $\lambda \geq 0$. Hence, from (a) and (b), $f(x^*) + \langle \lambda^*, g(x^*)\rangle \leq f(x^*) + \langle \lambda, g(x^*)\rangle$, $\forall \lambda \in \ell_+^{\infty,\star}$. (iii) ($\Leftarrow$) Suppose, for a given x^*, we have $\langle \lambda^*, g(x^*)\rangle = 0$, and $\lambda^* \in \arg\min_{\lambda \in \ell_+^{\infty,\star}} \mathcal{L}(x^*, \lambda)$. Suppose, for some r, $g_r(x^*) < 0$. Let $\hat{\lambda} = \lambda^* + \chi$, with $\chi(\nu) = 1$ if $\nu = r$ and 0 otherwise. Then $\mathcal{L}(x^*, \hat{\lambda}) < \mathcal{L}(x^*, \lambda^*)$, contradicting the fact that λ^* is a minimizer. Hence, it must be that $g(x^*) \geq 0$ (i.e., x^* is feasible). (iv) It immediately follows from (ii) and (iii) that if Assumption 7 holds and x^* solves (AP), then $\mathcal{L}$ has a saddle point in $\ell^\infty \times \ell_+^{\infty,\star}$. *Q.E.D.*

We now refine Theorem C.1 and give conditions such that the minimizing multiplier λ^* in Theorem C.1(ii) lies in $\ell^1 \subset \ell^{\infty,\star}$. By Yosida and Hewitt (1952), Theorems 1.22 and 1.24, $\ell^{\infty,\star}$ admits the decomposition $\ell^{\infty,\star} = \ell^1 + \ell^s$ with ℓ^s the set of *pure finitely additive* components. Assumption 8 below ensures summability of the minimizing multiplier.[40] In the assumption and throughout the proof, for a pair x and $y \in \ell^\infty$, and $T \in \mathbb{N}$, let $x^T(x, y) := x_r$ if $r \leq T$ and y_r, if $r > T$.

ASSUMPTION 8: (C) Continuity: $\lim_{T\to\infty} f(x^T(x, y)) = f(x)$.
(AN) Asymptotically non-anticipatory: $\forall t$, $\lim_{T\to\infty} g_r(x^T(x, y)) = g_r(x)$.
(AI) Asymptotically insensitive: for all N, $\lim_{r\to\infty}[g_r(x^N(x, y)) - g_r(y)] = 0$.
(B) Uniform boundedness: $\exists M$ s.t. for all T, $\sup_r \|g_r(x^T(x, y))\|_E \leq M$.[41]

THEOREM C.2: *Suppose f and g satisfy Assumptions 7(S) and 8. If $(x^*, \lambda^*) \in \ell^\infty \times \ell_+^{\infty,\star}$ is a saddle point of $\mathcal{L}$, then $(x^*, \lambda^*) \in \ell^\infty \times \ell_+^1$.*

PROOF OF THEOREM C.2: The proof uses two key lemmas.

LEMMA C.1: *Given Assumption 8(AI), $\forall \lambda^s \in \ell^s$ and $N \geq 1$, $\langle \lambda^s, g(x^N(x, y))\rangle = \langle \lambda^s, g(y)\rangle$.*

PROOF: If $\lambda^s \in \ell^s$, then for all $z = \{z_r\} \in \ell^\infty$ with $\lim_{r\to\infty} z_r = 0$, we have $\langle \lambda^s, z\rangle = 0$. By Assumption 8(AI), $\forall N$, $\lim_{r\to\infty}[g_r(x^N(x, y)) - g_r(y)] = 0$ and so $\forall N$, $\langle \lambda^s, [g(x^N(x, y)) - g(y)]\rangle = 0$ as required. *Q.E.D.*

LEMMA C.2: *Given Assumption 8(ANA) and (B), $\forall \lambda^1 \in \ell^1$, $\lim_{T\to\infty}\langle \lambda^1, g(x^T(x, y))\rangle = \langle \lambda^1, g(x)\rangle$.*

PROOF: For all $T, N \in \mathbb{N}$, $\|\langle \lambda^1, g(x^T(x, y))\rangle - \langle \lambda^1, g(x)\rangle\|_E \leq \sum_{r=0}^N \|\lambda_r^1\|_E \times \|g_r(x^T(x, y)) - g_r(x)\|_E + \sup_r \|g_r(x^T(x, y)) - g_r(x)\|_E \sum_{r=N+1}^\infty \|\lambda_t^1\|_E$. From Assumption 8(B), there is an $M > 0$ such that $\forall T$, $\sup_r \|g_r(x^T(x, y)) - g_r(x)\|_E \leq \bar{M} := M + \sup_r \|g_r(x)\|_E$. Since $\lambda^1 \in \ell^1$ for each $\varepsilon > 0$, there is an N_0 such that $\sum_{r=N_0+1}^\infty \|\lambda_r^1\|_E <$

[40] Dechert (1982) introduced the terminology used in Assumption 8(AN) and (AI). He showed summability of multipliers under slightly different assumptions. The proofs of Lemmas C.1 and C.2 follow Le Van and Sağlam (2004) who focused on variations to the deterministic model of optimal growth.

[41] The number $\|g_r(x)\|_E$ represents the Euclidean norm of the vector $g_r(x)$.

$\varepsilon/2\bar{M}$, and so, from Assumption 8(ANA), there is a $\bar{T}_r$ such that $\forall T \geq \bar{T}_r$, $\|\lambda_t^1\|_E \|g_r(x^T(x,y)) - g_r(x)\|_E < \varepsilon/2N_0$. Hence, combining conditions, for all $T > \max_{r\leq N_0}\{\bar{T}_r\}$, $\|\langle \lambda^1, g(x^T(x,y))\rangle - \langle \lambda^1, g(x)\rangle\|_E < N_0\, \varepsilon/2N_0 + \bar{M}\, \varepsilon/2\bar{M} = \varepsilon$. Since $\varepsilon > 0$ was arbitrary, this proves the result. *Q.E.D.*

We now conclude the proof of Theorem C.2. Since (x^*, λ^*) is a saddle in $\ell^\infty \times \ell_+^{\infty,\star}$, we have, for all $x \in \ell^\infty$, $\mathcal{L}(x, \lambda^*) \leq \mathcal{L}(x^*, \lambda^*)$, and $\langle \lambda^*, g(x^*)\rangle = 0$. Let $\lambda^* = \lambda^1 + \lambda^s$. Since $\lambda^* \geq 0$, we have both $\langle \lambda^1, g(x^*)\rangle = 0$ and $\langle \lambda^s, g(x^*)\rangle = 0$. Since x^* maximizes $\mathcal{L}(\cdot, \lambda^*)$ over ℓ^∞ and $\langle \lambda^*, g(x^*)\rangle = 0$,

$$f(x^*) \geq f\big(x^T(x^*, \hat{x})\big) + \big\langle \lambda^1, g\big(x^T(x^*, \hat{x})\big)\big\rangle + \big\langle \lambda^s, g\big(x^T(x^*, \hat{x})\big)\big\rangle. \tag{C.1}$$

Recall that if $\hat{x}$ is chosen to satisfy the Slater condition, then $\inf_r g_r(\hat{x}) > 0$. Lemmas C.1 and C.2 together imply that, for $T \to \infty$, we have both $\langle \lambda^1, g(x^T(x^*, \hat{x}))\rangle \to \langle \lambda^1, g(x^*)\rangle$ and $\langle \lambda^s, g(x^T(x^*, \hat{x}))\rangle \to \langle \lambda^s, g(\hat{x})\rangle$. Moreover, from Assumption 8(C), we have $f(x^T(x^*, \hat{x})) \to f(x^*)$. If $\lambda^s \neq 0$, since $\lambda^s \geq 0$, taking limits in (C.1), we have the following contradiction: $f(x^*) \geq f(x^*) + \langle \lambda^s, g(\hat{x})\rangle > f(x^*)$. Thus, $\lambda^s = 0$. *Q.E.D.*

Modified Problem. Let

$$\mathbf{P}^R = \big\{\mathbf{p} = (\mathbf{a}, \mathbf{k}, \mathbf{v}) |\ \forall t \geq 0,\ a_t : \mathcal{S}^t \to \mathcal{A},\ k_t : \mathcal{S}^{t-1} \to \mathcal{K},\ v_t^c : \mathcal{S}^t \to \mathcal{V}^c\big\}$$

denote the set of (modified) primal plans that exclude the quasilinear state variables. In addition, letting $b_0^l = \mathbb{I}_{n_l}$ (with $\mathbb{I}_{n_l}$ the n_l-identity matrix) and for all $\forall t, s^t, s' b_{t+1}^l(s^t, s'|s_0) = b_t^l(s^t|s_0) B^l(s_t, s')$, define

$$V^l\big(s_t, \mathbf{a}|s^t\big) := (1-\delta)\sum_{n=0}^{\infty} \delta^n \sum_{\mathcal{S}^n} b_n^l\big(s^{t+n}|s_t\big) u^l\big(s_{t+n}, a_{t+n}^l\big(s^{t+n}\big)\big) \pi^n\big(s^{t+n}|s_t\big), \tag{C.2}$$

where the previous expression is well defined since Assumption 5(iv) implies that all entries of the diagonal matrices B^l are bounded between -1 and $+1$. Relax the nonlinear laws of motion for state variables and replace the quasilinear state variables using (C.2) to obtain the modified problem

$$MP_0^* := \sup F\big[s_0, v_0^c, V^l(s_0, \mathbf{a})\big] \tag{MP}$$

subject to $\mathbf{p}^R \in \mathbf{P}^R$, $k_0 \leq \bar{k}$ and $\forall t, s^t$,

$$k_{t+1}\big(s^t\big) \leq W^k\big[k_t\big(s^{t-1}\big), s_t, a_t\big(s^t\big)\big], \tag{C.3}$$

$$v_t^c\big(s^t\big) \leq W^c\big[s_t, a_t\big(s^t\big), M^c\big[s_t, v_{t+1}^c\big(s^t, \cdot\big)\big]\big], \tag{C.4}$$

and

$$H\big[k_t\big(s^{t-1}\big), s_t, a_t\big(s^t\big), v_{t+1}^c\big(s^t, \cdot\big), V^l\big(\cdot, \mathbf{a}|s^t, \cdot\big)\big] \geq 0.$$

We relate the original problem (P) to the modified problem (MP).

LEMMA C.3: *Let Assumptions* 2, 4 *and* 5(iv) *hold and let* $\bar{k} \in \mathcal{K}^*(s_0)$. *Then* $MP_0^* = P_0^*$ *and for any solution* $\mathbf{p}^* = (\mathbf{p}^{R*}, \mathbf{v}^{l*})$ *to* (P), $\mathbf{p}^{R*}$ *is a solution to* (MP).

PROOF: Let $\mathbf{p} = (\mathbf{p}^R, \mathbf{v}^l)$ denote a feasible plan for (P); then, given Assumption 5(iv), it is readily shown via iteration on the law of motion for quasilinear states that $v_0 = V^l(s_0, \mathbf{a})$ and for all $t = 1, 2, \ldots$ and $s^t \in \mathcal{S}^t$, $v^l_t(s^t) = V^l(s_t, \mathbf{a}|s^t)$. Hence, since the nonlinear laws of motion are relaxed in (MP), $\mathbf{p}^R$ is feasible for (MP) and, so, $MP^*_0 \geq P^*_0$. Since $\bar{k} \in \mathcal{K}^*(s_0)$ and the constraints of (P) are non-empty, the constraint set for (MP) is also non-empty. By Assumption 2 and a similar argument to Proposition 1, (MP) has a solution $\hat{\mathbf{p}}^R$. Let $\hat{\mathbf{v}}^l = \{\hat{v}^l_t\}$, with $\hat{v}_0 = V^l(s_0, \hat{\mathbf{a}})$ and for all $t = 1, 2, \ldots$ and $s^t \in \mathcal{S}^t$, $\hat{v}^l_t(s^t) = V^l(s_t, \hat{\mathbf{a}}|s^t)$. If $\hat{\mathbf{p}} = (\hat{\mathbf{p}}^R, \hat{\mathbf{v}}^l)$ is feasible for (P), then $MP^*_0 = P^*_0$ and, hence, the $\mathbf{p}^R$ component of any solution to the original problem (P) solves (MP). Suppose that $\hat{\mathbf{p}}$ is not feasible for (P) and that $F[s_0, \hat{v}_0] > P^*_0$. Then $\hat{k}_0 \leq \bar{k}$,

$$\hat{k}_{t+1}(s^t) \leq W^k\big[\hat{k}_t(s^{t-1}), s_t, \hat{a}_t(s^t)\big], \tag{C.5}$$

$$\hat{v}^c_t(s^t) \leq W^c\big[s_t, \hat{a}_t(s^t), M^c\big[s_t, \hat{v}^c_{t+1}(s^t, \cdot)\big]\big], \tag{C.6}$$

$$\hat{v}^l_t(s^t) = W^l\big[s_t, \hat{a}_t(s^t), M^l\big[s_t, \hat{v}^l_{t+1}(s^t, \cdot)\big]\big],$$

and

$$H\big[\hat{k}_t(s^{t-1}), s_t, \hat{a}_t(s^t), \hat{v}_{t+1}(s^t, \cdot)\big] \geq 0,$$

with at least one of the constraints $\hat{k}_0 \leq \bar{k}$, (C.5), or (C.6) a strict inequality. Consider first modifying $\hat{\mathbf{p}}$ by increasing $\hat{k}_0$ until it equals $\bar{k} \in \mathcal{K}^*(s_0) \subset \mathcal{K}$ and successively at each history raising $\hat{k}_{t+1}(s^t)$ until it equals $W^k[\hat{k}_t(s^{t-1}), s_t, \hat{a}_t(s^t)]$. By Assumption 4(ii) and (iii), the modified plan satisfies (C.3) with equality at each s^t, the H constraints at each s^t, and has each $k_t(s^t) \in \mathcal{K}$. If each (C.6) holds with equality at $\hat{\mathbf{p}}$ and, hence, at the modified plan, then the modified plan has a payoff MP^*_0 (since it did not alter $\hat{v}_0$) and is feasible for (P). Thus, $P^*_0 \geq MP^*_0$. Suppose that at some s^t, $\hat{v}^c_t(s^t) < W^c[s_t, \hat{a}_t(s^t), M^c[s_t, \hat{v}^c_{t+1}(s^t, \cdot)]]$. Then further modify the plan by raising $\hat{v}^c_t(s^t)$ until equality is restored. Since we assume throughout that $W^v[s, a, M^v[s, \cdot]] : \mathcal{V}^{ns} \to \mathcal{V}$ and, hence, $W^c[s, a, M^c[s, \cdot]] : (\mathcal{V}^c)^{ns} \to \mathcal{V}^c$ and since $v^c_{t+1}(s^t, \cdot) \in (\mathcal{V}^c)^{ns}$, the adjusted $v^t_c(s^t) \in \mathcal{V}^c$. Continuing in this way through successively shorter histories s^τ, each $v^c_\tau(s^\tau)$ is increased (by the strict monotonicity of W^c and M^c) and, in particular, v^c_0 is increased. Hence, by the increasingness of $F[s_0, \cdot]$, the modified plan raises the value of $F[s_0, \cdot]$ above $F[s_0, \hat{v}_0]$. Since the $\mathbf{p}^R$ component of this plan is feasible for (MP), the optimality of $\hat{\mathbf{p}}^R$ for (MP) is contradicted. Thus, the modified plan must satisfy the conditions (C.4) with equality. We conclude that $P^*_0 \geq MP^*_0$. Combining inequalities $P^*_0 = MP^*_0$ and the $\mathbf{p}^R$ component of any optimum for (P) is feasible for (MP), attains a payoff of MP^*_0 and, hence, is optimal for (MP). *Q.E.D.*

Relating the Modified Problem to the Abstract Problem. For each (modified) primal plan $\mathbf{p}^R$, define the constraint values as follows: $\mathbf{z}(\mathbf{p}^R) = (\mathbf{z}^k(\mathbf{p}^R), \mathbf{z}^c(\mathbf{p}^R), \mathbf{z}^h(\mathbf{p}^R))$, with $\mathbf{z}^k(\mathbf{p}^R) = \{z^k_t(\mathbf{p}^R)\}_{t=0}^\infty$, where $z^k_0(\mathbf{p}^R) := \bar{k} - k_0$ and, for all $t = 1, 2, \ldots,$ and $s^t \in \mathcal{S}^t$,

$$z^k_t(\mathbf{p}^R)(s^t) := W^k\big[k_{t-1}(s^{t-1}), s_{t-1}, a_{t-1}(s^{t-1})\big] - k_t(s^t);$$

$\mathbf{z}^c(\mathbf{p}^R) = \{z^c_t(\mathbf{p}^R)\}_{t=0}^\infty$, where for all $t = 0, 1, 2, \ldots, s^t \in \mathcal{S}^t$,

$$z^c_t(\mathbf{p}^R)(s^t) := W^v\big[s_t, a_t(s^t), M^c\big[s_t, v^c_{t+1}(s^t, \cdot)\big]\big] - v^c_t(s^t)$$

and $\mathbf{z}^h(\mathbf{p}^R) = \{z^h_t(\mathbf{p}^R)\}_{t=0}^\infty$, where

$$z^h_t(\mathbf{p}^R)(s^t) := H\big[k_t(s^{t-1}), s_t, a_t(s^t), v^c_{t+1}(s^t, \cdot), V^l(s', \mathbf{a}|s^t, \cdot)\big].$$

Let $\mathbf{y}^k = \{y_t^k\}_{t=0}^{\infty}$, with $y_t^k : \mathcal{S}^t \to \mathbb{R}_+^{n_K}$, denote nonnegative costates for the backward-looking law of motion and $\mathbf{y}^c = \{y_t^c\}_{t=0}^{\infty}$, with $y_t^c : \mathcal{S}^t \to \mathbb{R}_+^{n_c}$, nonnegative costates for the forward-looking nonlinear laws of motion. Let $\mathbf{m} = \{m_t\}_{t=0}^{\infty}$, with $m_t : \mathcal{S}^t \to \mathbb{R}_+^{n_H}$, denote multipliers for the H-constraints. Collect these various multipliers into a (modified) *dual plan* $\mathbf{q}^R = \{\mathbf{m}, \mathbf{y}^k, \mathbf{y}^c\}$ and define the set of such dual plans to be

$$\mathbf{Q}^R = \left\{ \mathbf{q}^R \Big| \sum_{t=0}^{\infty} \sum_{\mathcal{S}^t} \delta^t \{ \| m_t(s^t) \|_E + \| y_t^k(s^t) \|_E + \| y_t^c(s^t) \|_E \} \pi^t(s^t | s_0) < \infty \right\}.$$

We associate the following Lagrangian $\mathcal{L} : \mathbf{P}^R \times \mathbf{Q}^R \to \mathbb{R}$ with (MP):

$$\mathcal{L}^R(\mathbf{p}^R, \mathbf{q}^R) = F[s_0, v_0^c, V^l(s_0, \mathbf{a}|s_0)] + \langle \mathbf{q}^R, \mathbf{z}(\mathbf{p}^R) \rangle,$$

with $\langle \mathbf{q}^R, \mathbf{z}(\mathbf{p}^R) \rangle = \sum_{t=0}^{\infty} \sum_{\mathcal{S}^t} \delta^t \{ m_t(s^t) \cdot z_t^h(\mathbf{p}^R)(s^t) + y_t^k(s^t) \cdot z_t^k(\mathbf{p}^R)(s^t) + y_t^c(s^t) \cdot z_t^c(\mathbf{p}^R) \times (s^t) \} \pi^t(s^t|s_0)$.

LEMMA C.4: *If $\mathbf{p}^{R*}$ solves (MP) and Assumptions 5 and 6 hold, then there exists a $\mathbf{q}^R \in \mathbf{Q}^R$ such that $(\mathbf{p}^{R*}, \mathbf{q}^{R*})$ is a saddle point of $\mathcal{L}^R$ on $\mathbf{P}^R \times \mathbf{Q}^R$.*

PROOF: We first re-express (MP) as an abstract problem of the form (AP). Next, we verify Assumptions 7 and 8. The result then follows from Theorems C.1 and C.2. In (MP) constraints are indexed by histories, in (AP) by the natural numbers. To convert one to the other, let $\mathcal{J} = \bigcup_{t=0}^{\infty} \mathcal{S}^t$ denote the countable set of all histories (and recall that $\mathcal{S} = \{1, \dots, n_S\}$). Relabel histories according to $\tau : \mathcal{J} \to \mathbb{N}$ with $\tau(s^t) = (1 + \sum_{r=1}^{t-1} s_r \cdot n_S^{t-r} + s_t)$.[42] Thus, the relabeled history 1 is the initial (date 0) history, relabeled histories $2, \dots, n_S + 1$ are the period 1 histories occurring after each period 1 shock realization, and so on. Choice variables may be grouped and re-indexed accordingly. Specifically, let $x = \{x_\tau\}_{\tau=1}^{\infty} \in \ell^\infty$ denote a regrouped and labeled primal plan, with

$$x_{\tau(s^t)} = (k_t(s^t), a_t(s^t), v_t^c(s^t)) \in \mathcal{X} = \mathcal{K} \times \mathcal{A} \times \mathcal{V}^{c, n_S},$$

where $\mathcal{V}^c$ denotes the bounded set of nonlinear states. Let g denote a regrouped and relabeled constraint function with $g = \{g_\tau\}_{\tau=1}^{\infty}$ and for each $\tau(s^t)$,

$$g_{\tau(s^t)}(x) = (z_t^k(\mathbf{p}^R(x))(s^t), z_t^c(\mathbf{p}^R(x))(s^t), z_t^h(\mathbf{p}^R(x))(s^t)),$$

where $\mathbf{p}^R(x) = \{p_t^R(x)(s^t)\}_{t=0}^{\infty}$ is the primal plan associated with x. The boundedness of the constraint functions implies $g : \ell^\infty \to \ell^\infty$. Finally, let $f(x) = F[s_0, v_0^c(x), V^l(s_0, \mathbf{a}(x))]$, where $v_0^c(x)$ and $\mathbf{a}(x)$ denote the v_0 and $\mathbf{a}$ components of x. In this way, (MP) is re-expressed in the form (AP) and, in particular, a Lagrangian of the form $\mathcal{L}(x, \lambda) = f(x) + \langle \lambda, g(x) \rangle$ may be associated with (MP).

We next show that Assumptions 5 and 6 in the main text and the structure of f and g in our setting imply Assumption 7 and 8. From Assumption 5, the concavity of F, H, W^c, and M^c imply concavity of f and g as required by Assumption 7(C). Assumption 6 ensures Assumption 7(S) (the Slater Condition). Specifically, if there is, as required by Assumption 6, a $\hat{\mathbf{p}} \in \mathbf{P}$ satisfying the law of motion constraints for $\mathbf{v}^c$ and $\mathbf{k}$ and the H

[42]For example, consider the history $(s_0, s_1, s_2) = (s_0, 2, 5)$. We have $\tau(s_0) = 1$, $\tau(s_0, 2) = 1 + 2 = 3$, and $\tau(s_0, 2, 5) = 1 + 2 \cdot n_S + 5 = 2 \cdot n_S + 6$.

constraints with strict inequality and $z^l(\hat{\mathbf{p}}) = 0$, then the corresponding $\hat{\mathbf{p}}^R \in \mathbf{P}^R$ satisfies Assumption 7(S).

Assumption 8(C) holds if for all x and y, $\lim_{T\to\infty} F[s_0, v_0^c(x^T(x,y)), V^l(s_0, \mathbf{a}(x^T(x,y)))] = F[s_0, v_0^c(x), V^l(s_0, \mathbf{a}(x))]$. Since $v_0^c(x^T(x,y)) = v_0^c(x)$ for all $T \geq 1$ and F is linear, so, continuous in its third argument, Assumption 8(C) holds if $\lim_{T\to\infty} V^l(s_0, \mathbf{a}(x^T(x,y))) = V^l(s_0, \mathbf{a}(x))$. For any pair x and y, there is a non-decreasing sequence of dates $R(T)$ with $\lim_{T\to\infty} R(T) = \infty$ such that, for all s^t with $t < R(T)$, all elements $p_t^R(x)(s^t) = p_t^R(x^T(x,y))(s^t)$. Thus,[43]

$$\begin{aligned}
&\left|V^l\big(s_0, \mathbf{a}\big(x^T(x,y)\big)\big) - V^l\big(s_0, \mathbf{a}(x)\big)\right| \\
&\quad \leq \delta^{R(T)} \sum_{s^{R(T)}} \left|b^l_{R(T)}\big(s^{R(T)}|s_0\big)\right| \\
&\qquad \times \left|V^l\big(s_{R(T)}, \mathbf{a}\big(x^T(x,y)\big)|s^{R(T)}\big) - V^l\big(s_{R(T)}, \mathbf{a}(x)|s^{R(T)}\big)\right| \pi^{R(T)}\big(s^{R(T)}|s_0\big).
\end{aligned}$$

Since $\delta^{R(T)}$ converges to 0 and the sequence of sums $\sum_{s^{R(T)}} |b^l_{R(T)}(s^{R(T)}|s_0)||V^l(s^t, \mathbf{a}(x^T(x,y))|s^t) - V^l(s^t, \mathbf{a}(x)|s^t)|\pi^{R(T)}(s^{R(T)}|s_0)$ is uniformly bounded, $\lim_{T\to\infty} |V^l(s_0, \mathbf{a}(x^T(x,y))) - V^l(s_0, \mathbf{a}(x))| = 0$. Thus, Assumption 8(C) is satisfied. Assumption 8(ANA) holds if for each $j = k, c, h$, and t and all x and y, $\lim_{T\to\infty} z_t^j(\mathbf{p}^R(x^T(x,y)))(s^t) = z_t^j(\mathbf{p}^R(x))(s^t)$. This result follows from the fact that each $z_t^j(\mathbf{p}^R)(s^t)$, $j = k, c$, depends only upon variables measurable with respect to s^t and s^{t+1} and each $z_t^h(\mathbf{p}^R)(s^t)$ depends only upon these variables and upon $V^l(s_t, \mathbf{a}|s^t)$ continuously. Thus, defining $R(T)$ as before, once $R(T) > t+1$, $z_t^j(\mathbf{p}^R(x^T(x,y))(s^t) = z_t^j(\mathbf{p}^R(x))(s^t)$, $j = k, c$. Also, by a similar argument to that given above, $|V^l(s_t, \mathbf{a}(x^T(x,y))|s^t) - V^l(s_t, \mathbf{a}(x)|s^t)| \to 0$. Thus, $\lim_{T\to\infty} z_t^h(\mathbf{p}^R(x^T(x,y))(s^t) = z_t^h(\mathbf{p}^R(x))(s^t)$ and Assumption 8(ANA) is verified.

Assumption 8(AI) requires that for each $j = c, k, h$, and N, $\lim_{t\to\infty}[z_t^j(\mathbf{p}^R(x^N(x,y)) - z_t^j(\mathbf{p}^R(y))] = 0$. This is an immediate consequence of the fact that each z_t^j does not depend on any variable that is measurable with respect to s^{t-1}. Thus, for any fixed N, there exists an $M(N)$ such that for each j, $t > M(N)$, and $r \geq 0$, $p^R_{t+r}(x^N(x,y))(s^{t+r}) = p^R_{t+r}(y)(s^{t+r})$ and so $[z_t^j(\mathbf{p}^R(x^N(x,y)) - z_t^j(\mathbf{p}^R(y))] = 0$. This confirms Assumption 8(AI). Finally, Assumption 8(B) follows from the boundedness of the constraint functions.

Thus, Assumptions 7 and 8 are verified. Let x^* denote the regrouped and relabeled primal plan corresponding to the solution $\mathbf{p}^{R*}$ of (MP). By Theorems C.1 and C.2, there exists a multiplier $\lambda^* \in \ell^1_+$ such that (x^*, λ^*) is a saddle point of $\mathcal{L}(x,\lambda) = f(x) + \langle \lambda, g(x)\rangle$, where f and g are those implied by (MP) and defined above. Resetting the labeling, it follows that $\mathbf{p}^{R*}$ and $\mathbf{q}^{R*}$ (where $\mathbf{q}^{R*}$ is obtained from λ^* by resetting the labeling and normalizing multipliers at each t and history s^t by $\delta^t \pi^t(s^t|s_0)$) is a saddle point of $\mathcal{L}^R$ on $\mathbf{P}^R \times \mathbf{Q}^R$. *Q.E.D.*

The preceding result establishes that if $\mathbf{p}^{R*}$ solves (MP), then there is a $\mathbf{q}^{R*} \in \mathbf{Q}^R$ such that $(\mathbf{p}^{R*}, \mathbf{q}^{R*})$ is a saddle point of $\mathcal{L}^R$ on $\mathbf{P}^R \times \mathbf{Q}^R$. We now seek to show that if $\mathbf{p}^*$ solves (P), then there is a $\mathbf{q}^* \in \mathbf{Q}$ such that $(\mathbf{p}^*, \mathbf{q}^*)$ is a saddle point of $\mathcal{L}$ on $\mathbf{P} \times \mathbf{Q}$. Recall that the Lagrangian $\mathcal{L}$ associated with (P) incorporates quasi-linear state variables and the laws of motion for such variables. It also allows the costates associated with nonlinear laws of motion to belong to $\mathbb{R}$ rather than $\mathbb{R}_+$. We use a constructive argument to show

[43]We use the following notation: for a generic matrix R, $|R|$ denotes the element-wise application of the absolute value operator to R.

that this Lagrangian also has a saddle point. Under Assumption 5, $F[s_0, v_0^c, v_0^l]$ is linear in its third argument. Below, we use the notation

$$F\big[s_0, v_0^c, v_0^l\big] = \hat{F}\big[s_0, v_0^c\big] + F_l(s_0) \cdot v_0^l.$$

In addition, $H[k, s, a, v^{c\prime}, v^{l\prime}]$ is linear in its final argument. Below, we use the notation

$$H\big[k, s, a, v^{c\prime}, v^{l\prime}\big] = \hat{H}\big[k, s, a, v^{c\prime}\big] + \delta \sum_{s' \in \mathcal{S}} N^l\big(s, s'\big) \cdot v^l\big(s'\big)\pi\big(s'|s\big).$$

LEMMA C.5: *Let* $\mathbf{p}^* = (\mathbf{p}^{R*}, \mathbf{v}^{l*})$, *with* $\mathbf{p}^{R*} = (\mathbf{a}^*, \mathbf{k}^*, \mathbf{v}^{c*})$, *be a solution to* (P). *If* $(\mathbf{p}^{R*}, \mathbf{q}^{R*})$, *with* $\mathbf{q}^{R*} = (\mathbf{y}^{k*}, \mathbf{y}^{c*}, \mathbf{m}^*)$, *is a saddle point of* $\mathcal{L}^R$ *on* $\mathbf{P}^R \times \mathbf{Q}^R$, *then* $(\mathbf{p}^{R*}, \mathbf{v}^{l*}, \mathbf{q}^{R*}, \mathbf{y}^{l*})$, *where* $\mathbf{y}^{l*}$ *satisfies the recursion* $y_0^{l*} = F^l(s_0)$ *and for all* $t = 1, 2, \dots$ *and* $s^t \in \mathcal{S}^t$,

$$y_{t+1}^{l*}\big(s^t, s'\big) = y_t^{l*}\big(s^t\big) \cdot B^l\big(s_t, s'\big) + m_t^*\big(s^t\big) \cdot N^l\big(s_t, s'\big) \tag{C.7}$$

is a saddle point of $\mathcal{L}$ *on* $\mathbf{P}^R \times \mathbf{Q}^R$.

PROOF: Let $\mathbf{Q}_+$ denote the subset of $\mathbf{Q}$ in which the costates $\mathbf{y}^k$ and $\mathbf{y}^c$ on backward-looking and nonlinear forward-looking state variables are nonnegative. For $(\mathbf{p}, \mathbf{q}) \in \mathbf{P} \times \mathbf{Q}_+$ and $(\mathbf{p}^R, \mathbf{q}^R) \in \mathbf{P}^R \times \mathbf{Q}^R$, define

$$\begin{aligned}
&\Delta\big(\mathbf{a}, \mathbf{v}^l, \mathbf{m}, \mathbf{y}^l\big) \\
&\quad := \mathcal{L}^R\big(\mathbf{p}^R, \mathbf{q}^R\big) - \mathcal{L}(\mathbf{p}, \mathbf{q}) \\
&\quad = F^l(s_0) \cdot V^l(s_0, \mathbf{a}) \\
&\qquad + \sum_{t=0}^{\infty} \delta^t \sum_{\mathcal{S}^t} m_t\big(s^t\big) \cdot \sum_{s_{t+1} \in \mathcal{S}} \delta N^l(s_t, s_{t+1}) V^l\big(s_{t+1}, \mathbf{a}|s^t, s_{t+1}\big)\pi(s_{t+1}|s_t)\pi^t\big(s^t|s_0\big) \\
&\qquad - F^l(s_0) \cdot v_0^l \\
&\qquad - \sum_{t=0}^{\infty} \delta^t \sum_{\mathcal{S}^t} m_t\big(s^t\big) \cdot \sum_{s_{t+1} \in \mathcal{S}} \delta N^l(s_t, s_{t+1}) v_{t+1}^l\big(s^t, s_{t+1}\big)\pi(s_{t+1}|s_t)\pi^t\big(s^t|s_0\big) \\
&\qquad - \sum_{t=0}^{\infty} \delta^t \sum_{\mathcal{S}^t} y_t^l\big(s^t\big) \cdot \Big[u^l\big(s_t, a_t\big(s^t\big)\big) + \delta \sum_{s_{t+1}} B^l(s_t, s_{t+1}) v_{t+1}^l\big(s^t, s_{t+1}\big)\pi(s_{t+1}|s_t) - v_t^l\big(s^t\big)\Big] \\
&\qquad \times \pi^t\big(s^t|s_0\big),
\end{aligned} \tag{C.8}$$

where the second equality follows from the definitions of the Lagrangians. From the recursion (C.7) defining the plan $\mathbf{y}^{l*}$ and the assumption $|B(s, s')| \le \mathbb{I}_{n_l}$, $\mathbf{y}^{l*}$ is summable with respect to the $\delta^t \pi^t(s^t|s_0)$ normalization as long as $\mathbf{m}^*$ is, that is, we have $\sum_{t=0}^{\infty} \sum_{s^t \in \mathcal{S}^t} \delta^t \|y_t^{l*}(s^t)\|_E \pi^t(s^t|s_0) < \infty$. Consequently, $(\mathbf{q}^{R*}, \mathbf{y}^{l*}) \in \mathbf{Q}_+$ and is feasible for the minimization defining the saddle point of $\mathcal{L}$. Hence, using the saddle point inequalities for $\mathcal{L}$ and $\mathcal{L}^R$ and the definition of Δ, $(\mathbf{p}^{R*}, \mathbf{v}^{l*}, \mathbf{q}^{R*}, \mathbf{y}^{l*})$ is a saddle point of $\mathcal{L}$ on $\mathbf{P} \times \mathbf{Q}_+$ if for all $(\mathbf{a}, \mathbf{v}^l) \in \mathbf{P}$ and $(\mathbf{m}, \mathbf{y}^l) \in \mathbf{Q}_+$,

$$\Delta\big(\mathbf{a}, \mathbf{v}^l, \mathbf{m}^*, \mathbf{y}^{l*}\big) \ge \Delta\big(\mathbf{a}^*, \mathbf{v}^{l*}, \mathbf{m}^*, \mathbf{y}^{l*}\big) \ge \Delta\big(\mathbf{a}^*, \mathbf{v}^{l*}, \mathbf{m}, \mathbf{y}^l\big).$$

Consider first $\Delta(\mathbf{a}^*, \mathbf{v}^{l*}, \mathbf{m}, \mathbf{y}^l)$. Since $(\mathbf{a}^*, \mathbf{v}^{l*})$ is part of an optimum and, hence, feasible for (P), it satisfies the law of motion for quasilinear states with equality. Consequently, the last line in (C.8) when evaluated at $(\mathbf{a}^*, \mathbf{v}^{l*})$ equals zero. Moreover, the law of motion for quasilinear states and Assumption 5(iv) imply that for all t, s^t, we have $v_t^*(s^t) = V^l(s_t, \mathbf{a}^*|s^t)$. Thus, (C.8) implies that for all $\mathbf{m}$ and $\mathbf{y}^l$ forming part of $\mathbf{q} \in \mathbf{Q}_+$,

$$\Delta\big(\mathbf{a}^*, \mathbf{v}^{l*}, \mathbf{m}, \mathbf{y}^l\big) = 0.$$

Consequently, the inequality $\Delta(\mathbf{a}^*, \mathbf{v}^{l*}, \mathbf{m}^*, \mathbf{y}^{l*}) \geq \Delta(\mathbf{a}^*, \mathbf{v}^{l*}, \mathbf{m}, \mathbf{y}^l)$ is trivially satisfied. Next, consider $\Delta(\mathbf{a}, \mathbf{v}^l, \mathbf{m}^*, \mathbf{y}^{l*})$. Substituting the recursion defining $\mathbf{y}^{l*}$ into (C.8) eliminates all terms involving $\mathbf{v}^l$ from $\Delta(\mathbf{a}, \mathbf{v}^l, \mathbf{m}^*, \mathbf{y}^{l*})$. Moreover, the definition of V^l implies that for all $\mathbf{a} \in \mathbf{A}$,

$$\begin{aligned}
&\sum_{t=0}^{\infty} \delta^t \sum_{\mathcal{S}^t} y_t^{l*}\big(s^t\big) \cdot u^l\big(s_t, a_t\big(s^t\big)\big) \\
&\quad = F^l(s_0) \cdot V^l(s_0, \mathbf{a}) \\
&\qquad + \sum_{t=0}^{\infty} \delta^t \sum_{\mathcal{S}^t} m_t^*\big(s^t\big) \cdot \sum_{s_{t+1} \in \mathcal{S}} \delta N^l(s_t, s_{t+1}) V^l\big(s_{t+1}, \mathbf{a}|s^t, s_{t+1}\big) \pi(s_{t+1}|s_t) \pi^t\big(s^t|s_0\big).
\end{aligned}$$

Substituting this into (C.8), it follows that $\Delta(\cdot, \cdot, \mathbf{m}^*, \mathbf{y}^{l*})$ is zero and independent of $\mathbf{a}$ and $\mathbf{v}^l$. Hence,

$$\Delta\big(\mathbf{a}, \mathbf{v}^l, \mathbf{m}^*, \mathbf{y}^{l*}\big) \geq \Delta\big(\mathbf{a}^*, \mathbf{v}^{l*}, \mathbf{m}^*, \mathbf{y}^{l*}\big), \quad \forall \mathbf{a}, \mathbf{v}^l \in \mathbf{P}.$$

Thus, $(\mathbf{p}^*, \mathbf{q}^*) = (\mathbf{p}^{R*}, \mathbf{v}^{l*}, \mathbf{q}^{R*}, \mathbf{y}^{l*})$ is a saddle point of $\mathcal{L}$ on $\mathbf{P} \times \mathbf{Q}_+$. It remains only to show that $(\mathbf{p}^{R*}, \mathbf{v}^{l*}, \mathbf{q}^{R*}, \mathbf{y}^{l*})$ is a saddle point of $\mathcal{L}$ on $\mathbf{P} \times \mathbf{Q}$ (i.e., of the Lagrangian without the costates $\mathbf{y}^k$ and $\mathbf{y}^c$ restricted to be nonnegative). However, since $\mathbf{p}^*$ is a solution to (P), it satisfies the laws of motion for backward-looking and nonlinear forward-looking states with equality. Hence, $\mathcal{L}(\mathbf{p}^*, \cdot)$ is independent of the multipliers on these states, implying that $\mathbf{q}^*$ minimizes $\mathcal{L}(\mathbf{p}^*, \cdot)$ on the set $\mathbf{Q}$. *Q.E.D.*

PROOF OF PROPOSITION 8: The proof now follows from the preceding lemmas. By Assumption 2, the restriction on $\bar{k}$, and Proposition 1, (P) has a solution $\mathbf{p}^* = (\mathbf{p}^{R*}, \mathbf{v}^{l*})$. Given Assumptions 2, 4, and 5(iv), by Lemma C.3, $\mathbf{p}^{R*}$ solves (MP). Given Assumptions 5 and 6, Lemma C.4 implies that the Lagrangian $\mathcal{L}^R$ has a saddle point $(\mathbf{p}^{R*}, \mathbf{q}^{R*})$. Finally, by Lemma C.5, the Lagrangian $\mathcal{L}$ has a saddle point. *Q.E.D.*

REFERENCES

ALIPRANTIS, C. D., AND K. C. BORDER (2006): *Infinite Dimensional Analysis: A Hitchhiker's Guide*. Berlin: Heidelberg. [2]

BERTSEKAS, D. P. (2003): *Nonlinear Programming* (Second Ed.). Belmont: Athena scientific. [3]

DECHERT (1982): "Lagrange Multipliers in Infinite Horizon Discrete Time Optimal Control Models," *Journal of Mathematical Economics*, 9, 285–302. [11]

FLORENZANO, M., AND C. L. VAN (2001): *Finite Dimensional Convexity and Optimization*, Vol. 13. Springer. [1]

JUDD, K., Ş. YELTEKIN, AND J. CONKLIN (2003): "Computing Supergame Equilibria," *Econometrica*, 71, 1239–1254. [1]

LE VAN, C., AND Ç. SAĞLAM (2004): "Optimal Growth Models and the Lagrange Multiplier," *Journal of Mathematical Economics*, 40, 393–410. [11]

LUENBERGER, D. (1969): *Optimization by Vector Space Methods*. New York: John Wiley & Sons. [11]

NECOARA, I., AND J. A. SUYKENS (2008): "Application of a Smoothing Technique to Decomposition in Convex Optimization," *Automatic Control, IEEE Transactions on*, 53 (11), 2674–2679. [3]
RUSZCZYŃSKI, A. P. (2006): *Nonlinear Optimization*, Vol. 13. Princeton University Press. [3]
SCHMUTZ, E. (2008): "Rational Points on the Unit Sphere," *Central European Journal of Mathematics*, 6 (3), 482–487. [2]
SLOAN, I. H., AND R. S. WOMERSLEY (2000): "Constructive Polynomial Approximation on the Sphere," *Journal of Approximation Theory*, 103 (1), 91–118. [1]
YOSIDA, K., AND E. HEWITT (1952): "Finitely Additive Measures," *Transactions of the Mathematical Society*, 72, 46–66. [11]

Co-editor Lars Peter Hansen handled this manuscript.

Manuscript received 16 September, 2013; final version accepted 24 May, 2017; available online 3 August, 2017.