



ECONOMIC RESEARCH
FEDERAL RESERVE BANK OF ST. LOUIS
WORKING PAPER SERIES

Extended Supplement to “Insurance and Inequality with Persistent Private Information”

Authors	Alex Bloedel, R. Vijay Krishna, and Oksana Leukhina
Working Paper Number	2025-012A
Creation Date	May 2025
Citable Link	https://doi.org/10.20955/wp.2025.012
Suggested Citation	Bloedel, A., Krishna, R.V., Leukhina, O., 2025; Extended Supplement to “Insurance and Inequality with Persistent Private Information”, Federal Reserve Bank of St. Louis Working Paper 2025-012. URL https://doi.org/10.20955/wp.2025.012

Federal Reserve Bank of St. Louis, Research Division, P.O. Box 442, St. Louis, MO 63166

The views expressed in this paper are those of the author(s) and do not necessarily reflect the views of the Federal Reserve System, the Board of Governors, or the regional Federal Reserve Banks. Federal Reserve Bank of St. Louis Working Papers are preliminary materials circulated to stimulate discussion and critical comment.

Extended Supplement to “Insurance and Inequality with Persistent Private Information”

Alexander W. Bloedel R. Vijay Krishna Oksana Leukhina

Abstract

This supplement contains auxiliary technical results and proofs omitted from Bloedel, Krishna, and Leukhina (2025) (henceforth BKL) and its Supplemental Appendix (henceforth SA). First, [Section I](#) proves parts (a)–(c) and (e)–(f) of [Theorem 3](#) from [Appendix B.1](#) of BKL. (Part (d) of [Theorem 3](#) is proved in SA-E.1 of BKL.) Second, [Section J](#) proves [Proposition 3.2](#) from [Section 3.3](#) of BKL and [Lemma B.2](#) from [Appendix B.2](#) of BKL. Finally, [Section K](#) proves supporting facts for Properties (a)–(e) from [Section 5.2](#) of BKL. Throughout, we follow the numbering and labeling conventions from BKL and its SA.

I. Proof of Theorem 3

This section provides the proof of [Theorem 3](#). It consists of several parts:

- [Section I.1](#) introduces the operator $\mathcal{B}_\circ$, of which D is the largest fixed point, and establishes basic existence and characterization results concerning D , D^* , and $\Gamma(\cdot)$.
- [Section I.2](#) presents the core of the proof, which involves analyzing two auxiliary operators, $\mathcal{B}$ and $\mathcal{B}^\dagger$. First, [Section I.2.1](#) establishes several results relating the fixed points of $\mathcal{B}_\circ$, $\mathcal{B}$, and $\mathcal{B}^\dagger$ to each other and to the set V_d . Second, [Sections I.2.2](#) and [I.2.3](#) characterize the iterates of these operators for the special case of CARA utility by analyzing systems of linear and conic programs. Third, [Section I.2.4](#) returns to the general case of DARA utility and shows that it reduces to the CARA case from [Section I.2.3](#).
- [Section I.3](#) constructs a useful suboptimal contract that is [TVC]-implementable and stays bounded away from the boundaries of D .
- [Section I.4](#) concludes by collecting the above pieces into a proof of [Theorem 3](#).

I.1. Preliminary Observations

We begin with an elementary lemma concerning the $\psi(\cdot, i, j)$ functions.

Lemma I.1. Let $i, j \in S$ with $i > j$. Under Assumption [DARA](#):

- (a) $\psi(u, i, j)$ is convex, $\psi'(u, i, j) \in (0, 1)$, and $\psi(u, i, j) > u$ for all $u \in \mathcal{U}$.
- (b) The extension of $\psi(\cdot, i, j)$ to $\mathbb{R}_-$ given by $\psi(0, i, j) := 0$ and $\psi'(0, i, j) := \lim_{u \rightarrow 0} \psi'(u, i, j)$ preserves continuity, convexity, and continuous differentiability.
- (c) $a\psi(u, i, j) \geq \psi(au, i, j)$ for all $u \in \mathcal{U}$ and $a \in [0, 1]$.

(d) Under CARA utility, $\psi(u, i, j) = (\theta_i/\theta_j) \cdot u$ where $\theta_k := e^{-\omega_k}$ for all $k \in S$. Thus, $a\psi(u, i, j) = \psi(au, i, j)$ for all $u \in \mathcal{U}$ and $a \in \mathbb{R}_+$.

Proof. For part (a), the Inverse Function Theorem yields

$$[\mathbf{I.1}] \quad \psi'(u, i, j) = \frac{U'(\omega_i + C(u, j))}{U'(\omega_j + C(u, j))} > 0$$

It follows directly from part (a) of Assumption **DARA** that $\psi(u, i, j) - u > 0$ and $\psi'(u, i, j) \in (0, 1)$. Moreover, we have

$$\log(\psi'(u, i, j)) = \log(U'(\omega_i + C(u, j))) - \log(U'(\omega_j + C(u, j)))$$

and because $C(\cdot, j)$ is strictly increasing, by part (c) of Assumption **DARA**, $\log(\psi'(u, i, j))$ is increasing. It follows that $\psi'(u, i, j)$ is increasing, i.e., $\psi(\cdot, i, j)$ is convex.

For part (b), note that, since the consumption domain is assumed unbounded above, $\lim_{u \rightarrow 0} \psi(u, i, j) = \lim_{u \rightarrow 0} U(\omega_i + C(u, j)) = \lim_{c \rightarrow \infty} U(\omega_i + c) = 0$. Thus this extension preserves continuity and convexity of $\psi(\cdot, i, j)$. Moreover, setting $\psi'(0, i, j) := \lim_{u \rightarrow 0} \psi'(u, i, j)$ preserves continuous differentiability by construction.

For part (c), let $u \in \mathcal{U}$ and $a \in [0, 1]$. We have $\psi(au, i, j) \leq a\psi(u, i, j) + (1 - a)\psi(0, i, j) = a\psi(u, i, j)$, where the inequality is from convexity and the equality is from $\psi(0, i, j) = 0$, shown above.

Part (d) is immediate from the definitions. $\square$

Recall that the set of *recursive constraints* consists of the of the *promise keeping* conditions

$$[\mathbf{PK}_j] \quad u_j + \alpha \mathbf{E}^{f_j} [\mathbf{w}_j] = v_j$$

for all $j \in S$, and the *incentive compatibility* conditions

$$[\mathbf{IC}_{ij}] \quad \psi(u_j, i, j) + \alpha \mathbf{E}^{f_i} [\mathbf{w}_j] \leq v_i$$

for all $i, j \in S$ with $i > j$.

For any $\mathbf{v} \in \mathcal{U}^d$, we say that the menu $(\mathbf{u}, \mathbf{w}) := (u_i, \mathbf{w}_i)_{i \in S} \in (\mathbb{R}_- \times \mathbb{R}^d)^d$ (where $\mathbf{u} := (u_1, \dots, u_d)$ and $\mathbf{w} := (\mathbf{w}_1, \dots, \mathbf{w}_d)$) *implements* $\mathbf{v}$ if it satisfies all of the recursive constraints at $\mathbf{v}$. (Note that implementation does not require even that $u_i < 0$ or $\mathbf{w}_i \ll \mathbf{0}$. Such feasibility requirements will be included as separate constraints.) Consider—in the manner of Abreu, Pearce, and Stacchetti (1990)—the operator $\mathcal{B}_\circ : 2^{\mathcal{U}^d} \rightarrow 2^{\mathcal{U}^d}$, defined

as follows for any $K \subset \mathcal{U}^d$:

$$[\text{I.2}] \quad \mathcal{B}_\circ(K) := \{\mathbf{v} \in \mathcal{U}^d : \exists (\mathbf{u}, \mathbf{w}) \text{ that implements } \mathbf{v} \text{ with } (u_i, \mathbf{w}_i) \in \mathcal{U} \times K \forall i \in S\}$$

It is easy to see that any fixed point of $\mathcal{B}_\circ$ is a recursive domain, and that the *largest recursive domain* D is precisely the *largest* fixed point of the operator $\mathcal{B}_\circ$.

Define the correspondence $\Gamma : \mathcal{U}^d \times \mathcal{K}_d \rightrightarrows (\mathcal{U} \times \mathcal{U}^d)^d$ by

$$[\text{I.3}] \quad \Gamma(\mathbf{v}, K) := \{(\mathbf{u}, \mathbf{w}) \in (\mathcal{U} \times K)^d : (\mathbf{u}, \mathbf{w}) \text{ implements } \mathbf{v}\}$$

Thus $\Gamma(\mathbf{v})$, defined in [3.1] in the main text, is equal to $\Gamma(\mathbf{v}, D)$. Clearly, an alternative way of defining the operator in [I.2] is $\mathcal{B}_\circ(K) = \{\mathbf{v} \in \mathcal{U}^d : \Gamma(\mathbf{v}, K) \neq \emptyset\}$.

Say that set $K \subseteq \mathbb{R}^d$ has *decreasing returns* if $\mathbf{v} \in K$ implies that $a\mathbf{v} \in K$ for all $a \in (0, 1]$. Say that a set $K \subseteq \mathbb{R}^d$ is *unbounded below* if, for all $r < 0$, there exists $\mathbf{v} \in K$ such that $\mathbf{v} \leq r\mathbf{1}$.

Lemma I.2. Let $K \subseteq \mathcal{U}^d$. Then:

- (a) If K is convex, then $\Gamma(\cdot, K)$ has convex graph.
- (b) If K has decreasing returns, then the graph of $\Gamma(\cdot, K)$ has decreasing returns.
- (c) If K is a cone and utility is CARA, then the graph of $\Gamma(\cdot, K)$ is a cone.

Proof. For part (a), suppose that K is convex and let $\gamma \in \Gamma(\mathbf{v}, K)$ and $\gamma' \in \Gamma(\mathbf{v}', K)$. Let $a \in (0, 1)$ and define $\mathbf{v}'' := a\mathbf{v} + (1-a)\mathbf{v}' \in \mathcal{U}^d$ and $\gamma'' := a\gamma + (1-a)\gamma' \in (\mathcal{U} \times K)^d$. Each of the [PK_j] constraints is linear, so clearly γ'' satisfies all of them at $\mathbf{v}''$. Consider constraint [IC_{ij}], with $i > j$. By definition of γ and γ' , we have

$$v_i'' \geq a\psi(u_j, i, j) + (1-a)\psi(u'_j, i, j) + \alpha \mathbf{E}^{\mathbf{f}_i} [a\mathbf{w}_j + (1-a)\mathbf{w}'_j]$$

and, since $\psi(\cdot, i, j)$ is convex by part (a) of Lemma I.1, $a\psi(u_j, i, j) + (1-a)\psi(u'_j, i, j) \geq \psi(u''_j, i, j)$. Thus, γ'' satisfies [IC_{ij}] at $\mathbf{v}''$. Hence, $\gamma'' \in \Gamma(\mathbf{v}'', K)$, i.e., $\Gamma(\cdot, K)$ has convex graph when K is convex.

For part (b), suppose that K has decreasing returns and let $\gamma \in \Gamma(\mathbf{v}, K)$. Let $a \in (0, 1)$ and define $\mathbf{v}' := a\mathbf{v} \in \mathcal{U}^d$ and $\gamma' := a\gamma \in (\mathcal{U} \times K)^d$. Each of the [PK_j] constraints is linear, so clearly γ' satisfies all of them at $\mathbf{v}'$. Consider constraint [IC_{ij}], with $i > j$. By definition of γ , we have

$$v'_i \geq a\psi(u_j, i, j) + \alpha \mathbf{E}^{\mathbf{f}_i} [a\mathbf{w}_j]$$

and, by part (c) of Lemma I.1, $a\psi(u_j, i, j) \geq \psi(au_j, i, j)$. Thus, γ' satisfies [IC_{ij}] at $\mathbf{v}'$. Hence, $\gamma' \in \Gamma(\mathbf{v}', K)$, i.e., $\Gamma(\cdot, K)$ has graph with decreasing returns when K has decreasing returns.

For part (c), mimic the proof of part (b) above, invoking part (d) of Lemma I.1

instead of part (c) of that lemma. $\square$

Proposition I.3. The following hold:

- (a) There exists a largest recursive domain D . It is nonempty, convex, has decreasing returns, and is unbounded below.
- (b) $D^* \subseteq D$ has all of the above properties.
- (c) The constraint correspondence $\Gamma : D \rightarrow (\mathcal{U} \times D)^d$ is nonempty-valued, has convex graph, and is continuous.

Proof. Begin with part (a). It is easy to see that the operator $\mathcal{B}_\circ$ is monotone increasing with respect to set inclusion. Because $(2^{\mathcal{U}^d}, \supseteq)$ defines a complete lattice, $\mathcal{B}_\circ$ has a largest fixed point D by Tarski's Fixed Point Theorem.

To establish that D is nonempty and unbounded below, define

$$B := \{c \in \mathbb{R} : c + \omega_i \in (\underline{c}, \infty) \ \forall i \in S\}$$

Clearly $\inf B = \underline{c} - \omega_1$ and $\sup B = +\infty$. For any $c \in B$, let $\mathbf{v}^c \in \mathcal{U}^d$ denote the vector of contingent promised utilities induced by the stochastic consumption process $(\omega^{(t)} + c)_{t=0}^\infty$. (Thus, $\mathbf{v}^0$ is the promised utility vector under autarky.) Define $D_c := \{\mathbf{v}^c\}$ for each $c \in B$. It is easy to see that $\mathcal{B}_\circ(D_c) = D_c$, and thus $\cup_{c \in B} D_c \subseteq D$, for each $c \in B$. Since $\mathcal{U} \subseteq \mathbb{R}_--$, it follows that

$$v_i^c \leq U(\omega_1 + c) \cdot T_1$$

for all $i \in S$, where

$$T_1 := \mathbf{E} \left[\sum_{t=0}^{\infty} \alpha^t \mathbf{1}(s^{(t+1)} = 1) \mid s^{(0)} = i \right]$$

and $T_1 > 0$ by Assumption [Markov](#). Moreover, $\lim_{c \rightarrow \inf B} U(\omega_1 + c) = -\infty$ under Assumption [DARA](#). Thus, for any $r \in \mathbb{R}_--$, there exists $c \in B$ such that $\max_{i \in S} v_i^c < r$. Thus, $\cup_{c \in B} D_c$ is unbounded below. Now, it follows from the definition of D that $D \supseteq \cup_{c \in B} D_c$. It follows that D is nonempty and unbounded below.

We now proceed to establish convexity and decreasing returns. Define $\mathcal{K}_d^{cdr} \subseteq 2^{\mathcal{U}^d}$ to be the collection of all subsets of $\mathcal{U}^d$ that are convex and have decreasing returns. It is easy to see that $\mathcal{K}_d^{cdr}$ is a complete lattice, with the meet operation defined by set intersection and the join operation defined by the natural notion of “convex, decreasing-returns hull,” namely

$$\bigvee_{\lambda \in \Lambda} K_\lambda := \bigcap \left\{ K \in \mathcal{K}_d^{cdr} : K \supseteq \bigcup_{\lambda \in \Lambda} K_\lambda \right\}$$

By Lemma [I.2](#), $\mathcal{B}_\circ : \mathcal{K}_d^{cdr} \rightarrow \mathcal{K}_d^{cdr}$ is well-defined and monotone increasing. Thus, by Tarski's Fixed Point Theorem, $\mathcal{B}_\circ$ has a largest fixed point $D' \in \mathcal{K}_d^{cdr}$. But since $\mathcal{U}^d \in \mathcal{K}_d^{cdr}$, it is easy to see that $D \in \mathcal{K}_d^{cdr}$ and hence $D = D'$ (see, e.g., the transfinite recursion in

Lemma 1 of Echenique, 2005). This concludes the proof of part (a).

Now consider part (b). Define scalar multiplication and convex combination of recursive contracts pointwise. It follows from the definition that if ξ satisfies [TVC] at $\mathbf{v} \in D$, then for any $a > 0$, $a\xi$ satisfies [TVC] at $a\mathbf{v} \in D$. Similarly, if ξ and ξ' satisfy [TVC] at, respectively, $\mathbf{v} \in D$ and $\mathbf{v}' \in D$, then for any $a \in (0, 1)$, $a\xi + (1 - a)\xi'$ satisfies [TVC] at $a\mathbf{v} + (1 - a)\mathbf{v}'$. It follows from these observations and the fact that D is convex and has decreasing returns, which was established in part (a), that D^* inherits both of these properties. Moreover, the vectors $\mathbf{v}^c$ defined in the proof of part (a) are clearly elements of D^* , which establishes that D^* is nonempty and unbounded below.

Finally, consider part (c). This follows immediately from part (a) of the present proposition and part (a) of Lemma I.2, namely the facts that D is convex and $\Gamma(\cdot) = \Gamma(\cdot, D)$ as defined in [I.3]. Continuity is easy to see from the definitions. $\square$

I.2. Fixed Points of Operators

While we are ultimately interested in (largest) fixed points of $\mathcal{B}_\circ$, it is convenient to work with some auxiliary operators. Consider first the operator $\mathcal{B}$, defined as follows for any $K \subset \mathbb{R}_-^d$:

$$[\text{I.4}] \quad \mathcal{B}(K) := \{ \mathbf{v} \in \mathbb{R}_-^d : \exists (\mathbf{u}, \mathbf{w}) \text{ that implements } \mathbf{v} \text{ with } (u_i, \mathbf{w}_i) \in \mathbb{R}_- \times K \ \forall i \in S \}$$

It is easy to see that $\mathcal{B}_\circ(K) \subset \mathcal{B}(K)$ for all K , so that the largest fixed point of $\mathcal{B}_\circ$ is a subset of the largest fixed point of $\mathcal{B}$.

Towards computing the largest fixed point of $\mathcal{B}$, it will be useful to consider the family of operators defined as follows for any $K \subset \mathcal{U}^d$:

$$[\text{I.5}] \quad \mathcal{B}_{ij}(K) := \{ \mathbf{v} \in \mathbb{R}_-^d : \exists (u_j, \mathbf{w}_j) \in \mathbb{R}_- \times K \text{ s.t. } [\mathbf{PK}_j] \text{ and } [\mathbf{IC}_{ij}] \text{ hold at } \mathbf{v} \}$$

We may use these $\mathcal{B}_{ij}$ operators to define the auxiliary operator $\mathcal{B}^\dagger : 2^{\mathbb{R}_-^d} \rightarrow 2^{\mathbb{R}_-^d}$ as

$$\mathcal{B}^\dagger(K) = \bigcap_{j=1}^{d-1} \mathcal{B}_{j+1,j}(K)$$

Since $\mathcal{B}^\dagger$ only accounts for local IC constraints, it is easy to see that $\mathcal{B}(K) \subset \mathcal{B}^\dagger(K)$ for all K , so that the largest fixed point of $\mathcal{B}$ is a subset of the largest fixed point of $\mathcal{B}^\dagger$.

We note the following property for future reference. Let $\mathcal{K} \subseteq 2^{\mathbb{R}_-^d}$ denote the collection of subsets $K \subseteq \mathbb{R}_-^d$ that are closed convex cones.

Lemma I.4. Suppose utility is CARA. The maps $\mathcal{B} : \mathcal{K} \rightarrow \mathcal{K}$ and $\mathcal{B}^\dagger : \mathcal{K} \rightarrow \mathcal{K}$ are well-defined.

Proof. Immediate from part (c) of Lemma I.2 and the definition of the constraint correspondence $\Gamma(\cdot)$ (see [3.1]). $\square$

I.2.1. Some Fixed Points

Let $V_d := \{\mathbf{v} \in \mathbb{R}_-^d : v_1 < v_2 < \dots < v_d\}$ and let $\text{cl } V_d$ denote its closure; thus, $\text{cl } V_d = \{\mathbf{v} \in \mathbb{R}_-^d : v_1 \leq v_2 \leq \dots \leq v_d\}$. We claim that, if the type process satisfies FOSD, $\mathcal{B}(\text{cl } V_d) = \text{cl } V_d$, $\mathcal{B}_o(\text{cl } V_d) = V_d$, and that $\mathcal{B}_o(V_d) = V_d$. We establish these facts below.

Lemma I.5. Consider $[\mathbf{PK}_j]$ and $[\mathbf{IC}_{ij}]$ where $i > j$ and let $\mathbf{v} \in \text{cl } V_d$.

- (a) Suppose the type process satisfies FOSD. Suppose, in addition, that $v_i = v_j$. Any solution to $[\mathbf{PK}_j]$ and $[\mathbf{IC}_{ij}]$ with $\mathbf{w}_j \in \text{cl } V_d$ must satisfy $u_j = 0$ and $\mathbf{E}^{\mathbf{f}_i}[\mathbf{w}_j] = \mathbf{E}^{\mathbf{f}_j}[\mathbf{w}_j]$.
- (b) In general, if $v_{j+1} > v_j$, there exist $u_j < 0$ and $\mathbf{w}_j \in \text{cl } V_d$ with $\mathbf{w}_j \ll \mathbf{0}$ such that $[\mathbf{PK}_j]$ holds and $[\mathbf{IC}_{ij}]$ holds as a strict inequality for all $i > j$. In particular, there exists such a solution in which $\mathbf{w}_j = w^j \mathbf{1}$ for some $w^j < 0$.

Proof. To see (a), let $(u_j, \mathbf{w}_j)$ satisfy both $[\mathbf{PK}_j]$ and $[\mathbf{IC}_{ij}]$. Subtract $[\mathbf{PK}_j]$ from each side of $[\mathbf{IC}_{ij}]$ to obtain

$$0 \leq \psi(u_j, i, j) - u_j + \alpha(\mathbf{E}^{\mathbf{f}_i}[\mathbf{w}_j] - \mathbf{E}^{\mathbf{f}_j}[\mathbf{w}_j]) \leq v_i - v_j = 0$$

Recall from Lemma I.1 that $\psi(u, i, j) > u$ for all $u \in \mathcal{U}$ and $\psi(0, i, j) = 0$. Moreover, $\mathbf{E}^{\mathbf{f}_i}[\mathbf{w}_j] \geq \mathbf{E}^{\mathbf{f}_j}[\mathbf{w}_j]$ because $\mathbf{w}_j \in \text{cl } V_d$ and $\mathbf{f}_i$ first-order stochastically dominates $\mathbf{f}_j$. Thus, the above display implies that $u_j = 0$ and $\mathbf{E}^{\mathbf{f}_i}[\mathbf{w}_j] = \mathbf{E}^{\mathbf{f}_j}[\mathbf{w}_j]$.

To see (b), notice that setting $u_j = 0$ and $\mathbf{w}_j = w^j \mathbf{1}$ with $w^j = v_j/\alpha$ means that $[\mathbf{PK}_j]$ holds as an equality, while $[\mathbf{IC}_{ij}]$ holds as a *strict* inequality for all $i > j$ (because $v_{j+1} > v_j$ and $\mathbf{v} \in \text{cl } V_d$). Now, increase w_j by a small amount while reducing u_j simultaneously to ensure that $[\mathbf{PK}_j]$ still holds. Because all the constraints $[\mathbf{IC}_{ij}]$ held as strict inequalities, this perturbation also satisfies these constraints as strict inequalities. $\square$

Lemma I.6. Fix $w^j < 0$. Then, for all $\varepsilon > 0$, there exists $\mathbf{w}_j \in V_d$ such that (i) $\mathbf{E}^{\mathbf{f}_j}[\mathbf{w}_j] = w^j$, and (ii) $\mathbf{E}^{\mathbf{f}_i}[\mathbf{w}_j] - \mathbf{E}^{\mathbf{f}_j}[\mathbf{w}_j] < \varepsilon$ for all $i > j$. If the type process satisfies FOSD, there exists $\mathbf{w}_j \in V_d$ that satisfies (i), (ii), and (iii) $\mathbf{E}^{\mathbf{f}_i}[\mathbf{w}_j] \geq \mathbf{E}^{\mathbf{f}_j}[\mathbf{w}_j]$ for all $i > j$.

Proof. Let $\varepsilon > 0$ be given. Let $\tilde{\varepsilon} > 0$ be a parameter to be optimised later. Let $w_{jk} := w^j - \tilde{\varepsilon}/2^k$ for $k = 1, \dots, d-1$, and $w_{jd} := w^j + \delta$, where $\delta := \frac{\tilde{\varepsilon}}{f_{jd}} \sum_{k=1}^{d-1} f_{jk}/2^k$. By construction, $w_{j1} < \dots < w_{jd}$. Thus, $\mathbf{w}_j \in V_d$. It is easy to see that $\mathbf{E}^{\mathbf{f}_j}[\mathbf{w}_j] = w^j$, establishing (i). To see (ii), note that $\mathbf{E}^{\mathbf{f}_i}[\mathbf{w}_j] - \mathbf{E}^{\mathbf{f}_j}[\mathbf{w}_j] \leq w_{jd} - w_{j1} = \delta + \tilde{\varepsilon}/2 \leq \tilde{\varepsilon}K$ for some $K > 0$. Thus, by choosing $\tilde{\varepsilon} \in (0, \varepsilon/K)$, we establish (ii). If the type process satisfies FOSD, it follows that $\mathbf{E}^{\mathbf{f}_i}[\mathbf{w}_j] \geq \mathbf{E}^{\mathbf{f}_j}[\mathbf{w}_j]$ for all $i > j$ because $w_{j1} < \dots < w_{jd}$. $\square$

Lemma I.7. The set V_d satisfies:

- (a) $\mathcal{B}(\text{cl } V_d) \supseteq \text{cl } V_d$.
- (b) $\mathcal{B}^\dagger(\text{cl } V_d) \supseteq \text{cl } V_d$.
- (c) $\mathcal{B}_\circ(V_d) \supseteq V_d$.

Proof. First, consider part (a). Let $\mathbf{v} \in \text{cl } V_d$. By part (b) of Lemma I.5, for each $j = 1, \dots, n-1$, there exist $(u_j, \mathbf{w}_j) \in \mathbb{R}_- \times \text{cl } V_d$ such that $[\mathbf{PK}_j]$ holds, and $[\mathbf{IC}_{ij}]$ holds for all $i > j$. Thus, $\text{cl } V_d \subseteq \mathcal{B}(\text{cl } V_d)$, which proves part (a). Part (b) follows immediately from part (a), as $\mathcal{B}^\dagger(K) \supseteq \mathcal{B}(K)$ for all $K \subseteq \mathbb{R}_-^d$.

Finally, consider part (c). Let $\mathbf{v} \in V_d$. Then, $v_j < v_{j+1}$ for all $j = 1, \dots, d-1$. By part (b) of Lemma I.5, there exist $u_j < 0$ and $\mathbf{w}_j \in \text{cl } V_d$ with $\mathbf{w}_j \ll \mathbf{0}$ such that $[\mathbf{PK}_j]$ holds and $[\mathbf{IC}_{ij}]$ holds as a strict inequality for all $i > j$. By Lemma I.6, we can replace $\mathbf{w}_j$ by some $\mathbf{w}'_j \in V_d$ so that $[\mathbf{PK}_j]$ still holds, and $[\mathbf{IC}_{ij}]$ holds as a strict inequality for all $i > j$. Because $\mathbf{w}'_j \in V_d$, it follows that $\mathcal{B}_\circ(V_d) \supseteq V_d$, as claimed. $\square$

Proposition I.8. The following hold:

- (a) There exists a largest fixed point K_* of $\mathcal{B}$. It satisfies $K_* \supseteq \text{cl } V_d$.
- (b) There exists a largest fixed point $K_*^\dagger$ of $\mathcal{B}^\dagger$. It satisfies $K_*^\dagger \supseteq K_* \supseteq \text{cl } V_d$.
- (c) The largest fixed point D of $\mathcal{B}_\circ$ satisfies $D \supseteq V_d$.

Proof. All three operators $\mathcal{B}$, $\mathcal{B}^\dagger$, and $\mathcal{B}_\circ$ are monotone increasing. Thus, by Tarski's Fixed Point Theorem, $\mathcal{B}$ and $\mathcal{B}^\dagger$ each has a largest fixed point in $\mathbb{R}_-^d$, K_* and $K_*^\dagger$, respectively. Existence of D , the largest fixed point of $\mathcal{B}_\circ$ in $\mathcal{U}^d$, was established in part (a) of Proposition I.3 by an identical argument.

Define $\Phi := \{K \subseteq \mathbb{R}_-^d : K \supseteq \text{cl } V_d\}$ and $\Phi_\circ := \{K \subseteq \mathcal{U}^d : K \supseteq V_d\}$. Both Φ and $\Phi_\circ$ define complete lattices when equipped with the set inclusion partial order. It follows from monotonicity and Lemma I.7 that the mappings $\mathcal{B} : \Phi \rightarrow \Phi$, $\mathcal{B}^\dagger : \Phi \rightarrow \Phi$, and $\mathcal{B}_\circ : \Phi_\circ \rightarrow \Phi_\circ$ are well-defined. Thus, by Tarski's Fixed Point Theorem, $\mathcal{B}$ and $\mathcal{B}^\dagger$ each has a largest fixed point in Φ and $\mathcal{B}_\circ$ has a largest fixed point in $\Phi_\circ$. It follows immediately that $K_*, K_*^\dagger \in \Phi$ and $D \in \Phi_\circ$. $\square$

Proposition I.9. Suppose that the type process satisfies FOSD. Then:

- (a) $\mathcal{B}(\text{cl } V_d) = \text{cl } V_d$.
- (b) $\mathcal{B}^\dagger(\text{cl } V_d) = \text{cl } V_d$.
- (c) $\mathcal{B}_\circ(\text{cl } V_d) \subseteq V_d \subsetneq \text{cl } V_d$.

Proof. First, consider part (a). By part (a) of Lemma I.7, it suffices to show that $\mathcal{B}(\text{cl } V_d) \subset \text{cl } V_d$ under the additional FOSD assumption. To that end, let $\mathbf{v} \in \mathbb{R}_-^d$ and consider $i > j$. Then, subtracting $[\mathbf{PK}_j]$ from $[\mathbf{IC}_{ij}]$, we find that

$$0 \leq \psi(u_j, i, j) + \alpha(\mathbf{E}^{f_i}[\mathbf{w}_j] - \mathbf{E}^{f_j}[\mathbf{w}_j]) \leq v_i - v_j$$

which implies $v_i \geq v_j$, as (i) $\psi(u_j, i, j) - u_j \geq 0$, and (ii) $\mathbf{E}^{f_i}[\mathbf{w}_j] \geq \mathbf{E}^{f_j}[\mathbf{w}_j]$ because f_i first order stochastically dominates f_j and $\mathbf{w}_j \in \text{cl } V_d$. Thus, if $\mathbf{v} \in \mathcal{B}(\text{cl } V_d)$, we must have $v_i \geq v_j$ whenever $i > j$, i.e., $\mathcal{B}(\text{cl } V_d) \subset \text{cl } V_d$.

Next, consider part (b). By part (b) of Lemma I.7, it suffices to show that $\mathcal{B}^\dagger(\text{cl } V_d) \subset \text{cl } V_d$ under the additional FOSD assumption. The proof of this fact is identical to the proof for part (a).

Finally, consider part (c). It is immediate that $\mathcal{B}_\circ(\text{cl } V_d) \subseteq \mathcal{B}(\text{cl } V_d) = \text{cl } V_d$ by monotonicity. Take any $\mathbf{v} \in \text{cl } V_d \setminus V_d$. It is easy to see that there exists some $j \in S$ such that either (i) $v_j = 0$ or (ii) $v_j = v_{j+1}$ or both. In case (i), [PK_j] implies that

$$0 = u_j + \alpha \mathbf{E}^{f_j}[\mathbf{w}_j] \leq u_j$$

where the inequality follows from $\mathbf{w}_j \in \text{cl } V_d$. It follows that, in case (i), it must be that $u_j = 0$ and hence $\mathbf{v} \notin \mathcal{B}_\circ(\text{cl } V_d)$. In case (ii), by part (a) of Lemma I.5, if $(u_j, \mathbf{w}_j) \in \mathbb{R}_- \times \text{cl } V_d$ satisfy [IC_{ij}] and [PK_j], it must necessarily be the case that $u_j = 0$. Again, it follows that $\mathbf{v} \notin \mathcal{B}_\circ(\text{cl } V_d)$, which completes the proof of the claim. $\square$

Proposition I.10. If the type process satisfies FOSD, then V_d is a fixed point of $\mathcal{B}_\circ$.

Proof. Lemma I.7 states that $\mathcal{B}_\circ(V_d) \supseteq V_d$. Part (c) of Proposition I.9 states that $\mathcal{B}_\circ(\text{cl } V_d) \subseteq V_d$ and, since the operator $\mathcal{B}_\circ$ is monotone, it follows that $\mathcal{B}_\circ(V_d) \subseteq V_d$. This completes the proof. $\square$

I.2.2. Iterates of $\mathcal{B}^\dagger$ in the CARA Case

Recall (from, e.g., part (c) of Lemma I.1) that under CARA utility, $\psi(u, i, j) = (\theta_i/\theta_j) \cdot u$ where $\theta_i := e^{-\omega_i}$. Thus, the incentive constraints [IC_{ij}] are linear in the u_j .

For any $\mathbf{t} = (t_2, \dots, t_d) \in \mathbb{R}_+^{d-1}$, let

$$[\text{I.6}] \quad K_{\mathbf{t}} := \{\mathbf{v} \in \mathbb{R}_-^d : v_{j-1}t_j \leq v_j, \text{ for } j = 2, \dots, d\}$$

It is easy to see that $K_{\mathbf{t}}$ is a closed, convex cone and is a subset of $\mathbb{R}_-^d$.

Proposition I.11. Suppose utility is CARA. There exists $\mathbf{t} = (t_2, \dots, t_d) \geq \mathbf{1}$ with $t_2 \cdot t_3 \cdots t_d < \infty$ such that $\mathcal{B}^\dagger(\mathbb{R}_-^d) = K_{\mathbf{t}}$.

Proof. Let us first consider $\mathcal{B}_{21}(\mathbb{R}_-^d)$. We want to characterize the set of $\mathbf{v} \in \mathbb{R}_-^d$ that can be implemented while satisfying

$$[\text{PK}_1] \quad u_1 + \alpha \mathbf{E}^{f_1}[\mathbf{w}_1] = v_1$$

and

$$[\mathbf{IC}_{21}] \quad \frac{\theta_2}{\theta_1} \cdot u_1 + \alpha \mathbf{E}^{f_2}[\mathbf{w}_1] \leq v_2$$

subject to

$$[\mathbf{Feas}_{21}] \quad u_1 \leq 0 \quad \text{and} \quad \mathbf{w}_1 \in \mathbb{R}_-^d$$

It is easy to see that any $\mathbf{v}$ with $v_2 \geq v_1$ can be implemented. (Indeed, this follows from part (b) of Lemma I.7.) For general $\mathbf{v}$, consider the following problem: What is the maximum value of $u_1 + \alpha \mathbf{E}^{f_1}[\mathbf{w}_1]$ that can be achieved while satisfying $[\mathbf{IC}_{21}]$ and the feasibility constraints $[\mathbf{Feas}_{21}]$?

Our question can be formulated as the following linear programming problem

$$\begin{aligned} \max_{u_1, \mathbf{w}_1} \quad & u_1 + \alpha \sum_{j=1}^d f_{1j} w_{1j} \\ \text{s.t.} \quad & [\mathbf{IC}_{21}] \text{ and } [\mathbf{Feas}_{21}] \end{aligned}$$

It is easy to see that the solution to this problem consists of setting $u_1 = 0$, letting k^* such that $f_{1k^*}/f_{2k^*} \leq f_{1j}/f_{2j}$ for all $j = 1, \dots, d$, and setting $w_{1j} = 0$ for $j \neq k^*$ while choosing w_{1k^*} so that $[\mathbf{IC}_{21}]$ holds with equality. Intuitively, $[\mathbf{IC}_{21}]$ is a budget constraint and our solution consists of choosing the smallest amount of the ‘quantity’ (u_1 and $\mathbf{w}_1$) where the marginal utility-to-price ratio is smallest, while choosing the largest quantity of everything else.

This implies that the largest value $u_1 + \alpha \sum_{j=1}^d f_{1j} w_{1j}$, and hence v_1 , can take is $v_2 \cdot f_{1k^*}/f_{2k^*}$. That is, $\mathcal{B}_{21}(\mathbb{R}_-^d) := \{\mathbf{v} \in \mathbb{R}_-^d : v_2 \geq t_2 v_1, t_2 := \max[f_{2j}/f_{1j}]\}$. Similarly, we can show that $\mathcal{B}_{j+1,j}(\mathbb{R}_-^d) := \{\mathbf{v} \in \mathbb{R}_-^d : v_{j+1} \geq t_{j+1} v_j, t_{j+1} := \max[f_{j+1,k}/f_{jk}]\}$. Clearly, $1 \leq t_j < \infty$ for every $j = 2, \dots, d$, which proves the claim. $\square$

We will use the following version of Linear Programming Duality and Complementary Slackness in the sequel.¹

Theorem 4 (LP duality and Complementary Slackness). *Let $\mathbf{y}^* = (y_1^*, \dots, y_n^*)$ be a feasible solution of the linear program*

$$[\mathbf{P}] \quad \max_{\mathbf{y}} \quad c \cdot \mathbf{y} \quad \text{subject to } A\mathbf{y} \leq \mathbf{b} \text{ and } \mathbf{y} \leq \mathbf{0}$$

¹Section 4.3 of Vohra (2005) shows that $[\mathbf{P}]$ and $[\mathbf{D}]$ as written here are dual to each other. Alternatively, see Section 6.2 of Matoušek and Gärtner (2007). The version of complementary slackness described in Theorem 4 is on p. 204 of the Glossary of Matoušek and Gärtner (2007).

and let $q^* = (q_1^*, \dots, q_m^*)$ be a feasible solution of the dual linear program

$$[\mathbf{D}] \quad \min_q \quad b \cdot q \quad \text{subject to } qA \leq c \text{ and } q \geq 0$$

Then, the following are equivalent:

- (a) y^* is optimal for $[\mathbf{P}]$ and q^* is optimal for $[\mathbf{D}]$.
- (b) For all $i = 1, \dots, m$, y^* satisfies the i th constraint of $[\mathbf{P}]$ with equality or $q_i^* = 0$; similarly, for all $j = 1, \dots, n$, q^* satisfies the j th constraint of $[\mathbf{D}]$ with equality or $y_j^* = 0$.

Consider the operator $\mathcal{B}_{j+1,j}(K_t)$ where $t \geq 1$ is finite in every component. We seek to characterize the set of all $\mathbf{v}$ that can be implemented by $(u_j, \mathbf{w}_j)_{j \in S} \in \mathbb{R}_-^d \times K_t$ while satisfying $[\mathbf{PK}_j]$ and $[\mathbf{IC}_{ij}]$ where $i = j + 1$. As in the proof of Proposition I.11, for a given v_{j+1} , we wish to find the largest value of v_j that can be implemented. We can state this problem as follows:

$$\begin{aligned}
 [\mathbf{P}_{j+1,j}] \quad & \max_{u_j, \mathbf{w}_j} \quad u_j + \alpha[f_{j1}w_{j1} + f_{j2}w_{j2} + \dots + f_{jd}w_{jd}] \\
 \text{s.t.} \quad & \frac{\theta_{j+1}}{\theta_j} \cdot u_j + \alpha[f_{j+1,1}w_{j,1} + \dots + f_{j+1,d}w_{j,d}] \leq v_{j+1} \\
 & \alpha t_2 w_{j1} - \alpha w_{j2} \leq 0 \\
 & \alpha t_3 w_{j2} - \alpha w_{j3} \leq 0 \\
 & \vdots \\
 & \alpha t_d w_{j,d-1} - \alpha w_{jd} \leq 0 \\
 & u_1, w_{jd} \leq 0
 \end{aligned}$$

where we note that $u_j + \alpha[f_{j1}w_{j1} + f_{j2}w_{j2} + \dots + f_{jd}w_{jd}] = v_j$. It is easy to see that $[\mathbf{P}_{j+1,j}]$ has a solution. Let $\mathbf{q}^{j+1}$ be a feasible solution of the dual linear program

$$\begin{aligned}
 [\mathbf{D}_{j+1,j}] \quad & \min_{\mathbf{q}^{j+1}} \quad q_1^{j+1} \cdot v_{j+1} \\
 \text{s.t.} \quad & \theta_{j+1} q_1^{j+1} \leq \theta_j \\
 & \alpha f_{j+1,1} q_1^{j+1} + \alpha t_2 q_2^{j+1} \leq \alpha f_{j1} \\
 & \alpha f_{j+1,2} q_1^{j+1} - \alpha q_2^{j+1} + \alpha t_3 q_3^{j+1} \leq \alpha f_{j2} \\
 & \vdots \\
 & \alpha f_{j+1,d-1} q_1^{j+1} - \alpha q_{d-1}^{j+1} + \alpha t_d q_d^{j+1} \leq \alpha f_{j,d-1} \\
 & \alpha f_{j+1,d} q_1^{j+1} - \alpha q_d^{j+1} \leq \alpha f_{jd} \\
 & \mathbf{0} \leq \mathbf{q}^{j+1}
 \end{aligned}$$

Note that, for any $v_{j+1} \leq 0$, solving $[\mathbf{D}_{j+1,j}]$ is equivalent to maximizing q_1^{j+1} subject to the above constraints. The dual constraints can be rewritten as follows:

$$\begin{aligned}
[\mathbf{DC}_0] \quad & q_1^{j+1} \leq \frac{\theta_j}{\theta_{j+1}} \\
[\mathbf{DC}_1] \quad & q_1^{j+1} \leq \frac{f_{j1}}{f_{j+1,1}} - \frac{t_2}{f_{j+1,1}} q_2^{j+1} \\
& \vdots \\
[\mathbf{DC}_k] \quad & q_1^{j+1} \leq \frac{f_{jk}}{f_{j+1,k}} + \frac{q_k^{j+1}}{f_{j+1,k}} - \frac{t_{k+1}}{f_{j+1,k}} q_{k+1}^{j+1} \\
& \vdots \\
[\mathbf{DC}_{d-1}] \quad & q_1^{j+1} \leq \frac{f_{j,d-1}}{f_{j+1,d-1}} + \frac{q_{d-1}^{j+1}}{f_{j+1,d-1}} - \frac{t_d}{f_{j+1,d-1}} q_d^{j+1} \\
[\mathbf{DC}_d] \quad & q_1^{j+1} \leq \frac{f_{jd}}{f_{j+1,d}} + \frac{q_d^{j+1}}{f_{j+1,d}}
\end{aligned}$$

where $q_1^{j+1} \geq 0$ and $[\mathbf{DC}_k]$ is the k th dual constraint.

Define inductively the sequence $\mathbf{t}^{(n)} \in \mathbb{R}_+^{d-1}$ as follows: $K_{\mathbf{t}^{(1)}} := \mathcal{B}^\dagger(\mathbb{R}_-^d)$, and for all $n > 1$, $K_{\mathbf{t}^{(n)}} := \mathcal{B}^\dagger(K_{\mathbf{t}^{(n-1)}})$.

Lemma I.12. The sequence $(\mathbf{t}^{(n)}) \subseteq \mathbb{R}_+^{d-1}$ is well-defined and satisfies:

- (a) $\mathbf{t}^{(n)} \geq \mathbf{1}$ for all $n \geq 1$.
- (b) $\mathbf{t}^{(n)} \leq \mathbf{t}^{(n-1)}$ for all $n \geq 1$.
- (c) The largest fixed point of $\mathcal{B}^\dagger$, $K_*^\dagger = \lim_{n \rightarrow \infty} K_{\mathbf{t}^{(n)}}$.
- (d) For each $j = 1, \dots, d-1$, $t_{j+1}^{(n)} = 1/q_1^{j+1}(\mathbf{t}^{(n-1)})$ where $q_1^{j+1}(\mathbf{t}^{(n-1)})$ is the solution to $[\mathbf{D}_{j+1,j}]$ given $\mathbf{t}^{(n-1)}$.

Proof. Proposition I.11 shows that $\mathbf{t}^{(1)}$ is well-defined. The above linear programming formulation shows that $\mathbf{t}^{(n)}$ is well-defined for all $n > 1$ and is characterized as stated in part (d) of the present lemma. It remains to show parts (a)–(c). Part (b) follows from monotonicity of $\mathcal{B}^\dagger$. Part (c) follows from the simple observation that $\mathcal{B}^\dagger$ is order-continuous on the complete lattice $(2^{\mathbb{R}_-^d}, \supseteq)$ and arguments from Section 4 of Echenique (2005). Finally, suppose part (a) fails, i.e., there exists some $n \geq 1$ such that $\mathbf{t}^{(n)} \not\geq \mathbf{1}$. Then, by parts (b) and (c), it follows that $K_*^\dagger \not\supseteq V_d$, which contradicts part (b) of Proposition I.8. $\square$

Corollary I.13. Given any $j = 1, \dots, d-1$ and $\mathbf{t}^{(n)}$ for any $n \geq 1$, constraint $[\mathbf{DC}_0]$ holds as a strict inequality at any optimal solution of $[\mathbf{D}_{j+1,j}]$. Thus, $u_j = 0$ at any optimal solution of $[\mathbf{P}_{j+1,j}]$.

Proof. Towards a contradiction, suppose $[\mathbf{DC}_0]$ holds as an equality at some optimal solution given $\mathbf{t}^{(n)}$. This implies $q_1^{j+1} = \theta_j/\theta_{j+1} > 1$. Then, by part (d) of Lemma I.12,

$t_{j+1}^{(n+1)} < 1$. But this contradicts part (a) of Lemma I.12. That $u_j = 0$ at any optimum then follows from Complementary Slackness (part (b) of Theorem 4). $\square$

The following lemma describes which constraints bind in the dual program. In what follows, we say that a constraint *binds* if it holds with equality and is *slack* if it holds as a strict inequality.

Lemma I.14. Suppose the type process satisfies MLRP and $f_j \neq f_{j+1}$. Then, given any $t \geq 1$, there exists an optimal solution $\mathbf{q}^{j+1}$ to $[\mathbf{D}_{j+1,j}]$ where $q_k^{j+1} > 0$ and $[\mathbf{DC}_k]$ binds for all $k = 1, \dots, d$.

Proof. To begin, note that MLRP and $f_j \neq f_{j+1}$ imply that $k \mapsto \frac{f_{jk}}{f_{j+1,k}}$ is non-increasing and $\frac{f_{j1}}{f_{j+1,1}} > 1 > \frac{f_{jd}}{f_{j+1,d}}$. Fix any optimal solution $\mathbf{q}^{j+1} = (q_1^{j+1}, \dots, q_d^{j+1})$ of $[\mathbf{D}_{j+1,j}]$ such that $q_1^{j+1} \geq \tilde{q}_1^{j+1}$ for all optimal solutions $\tilde{\mathbf{q}}^{j+1}$ of $[\mathbf{D}_{j+1,j}]$.² We proceed by induction.

For the base step, we first claim that $[\mathbf{DC}_1]$ binds. Suppose, towards a contradiction, that this constraint is slack. Then there exists an $\varepsilon > 0$ such that, holding q_1^{j+1} fixed, setting $\hat{q}_2^{j+1} := q_2^{j+1} + \frac{f_{j+1,1}}{t_2} \frac{\varepsilon}{2}$ and (if $d \geq 3$) inductively setting $\hat{q}_{k+1}^{j+1} := q_{k+1}^{j+1} + \frac{f_{j+1,k}}{t_{k+1}} \left(\frac{\hat{q}_k^{j+1} - q_k^{j+1}}{2f_{j+1,k}} \right)$ for all $k = 2, \dots, d-1$ yields a feasible solution $(q_1^{j+1}, \hat{q}_2^{j+1}, \dots, \hat{q}_d^{j+1})$ at which all of the $[\mathbf{DC}_k]$ constraints are slack. Thus, there must exist a minimiser of $[\mathbf{D}_{j+1,j}]$ that strictly exceeds q_1^{j+1} , delivering the desired contradiction. Therefore, $[\mathbf{DC}_1]$ binds at $\mathbf{q}^{j+1}$. This implies that $q_2^{j+1} > 0$, because Lemma I.12 implies that $q_1^{j+1} \leq 1$ and, as noted above, $\frac{f_{j1}}{f_{j+1,1}} > 1$.

For the inductive step, let $1 \leq \ell < d-1$ be given and suppose that $[\mathbf{DC}_k]$ binds and $q_{k+1}^{j+1} > 0$ for all $k = 1, \dots, \ell$. Suppose, towards a contradiction, that $[\mathbf{DC}_k]$ for $k = \ell+1$ is slack. This implies that we can: (i) slightly reduce each $q_k^{j+1} > 0$ for $2 \leq k \leq \ell+1$ so that all $[\mathbf{DC}_k]$ for $1 \leq k \leq \ell+1$ become slack, and then (ii) slightly increase each $q_k^{j+1} \geq 0$ for $k > \ell+1$ so that all $[\mathbf{DC}_k]$ for $k > \ell+1$ become slack as well (as in the base step). Thus, there must exist a minimiser of $[\mathbf{D}_{j+1,j}]$ that strictly exceeds q_1^{j+1} , delivering the desired contradiction. Therefore $[\mathbf{DC}_k]$ for $k = \ell+1$ must bind. Suppose, towards a contradiction,

²Note that this selection of optimal solutions is only relevant if $v_{j+1} = 0$; in particular, if $v_{j+1} < 0$, then all optimal solutions must specify the same value of q_1^{j+1} .

that $q_{\ell+1}^{j+1} = 0$. Then we have

$$\begin{aligned}
q_1^{j+1} &= \frac{f_{j,\ell+1}}{f_{j+1,\ell+1}} - \frac{t_{\ell+2}}{f_{j+1,\ell+1}} q_{\ell+2}^{j+1} \\
&\leq \frac{f_{j,\ell+1}}{f_{j+1,\ell+1}} \\
&\leq \frac{f_{j,\ell}}{f_{j+1,\ell}} \\
&= q_1^{j+1} - \frac{q_\ell^{j+1}}{f_{j+1,\ell}} \\
&< q_1^{j+1}
\end{aligned}$$

where the first line is by the fact (shown above) that $[\mathbf{DC}_k]$ for $k = \ell + 1$ binds and the supposition that $q_{\ell+1}^{j+1} = 0$, the second line is by $q^{j+1} \geq 0$, the third line is by MLRP, and the final two lines are by the inductive hypotheses that $[\mathbf{DC}_k]$ for $k = \ell$ binds and $q_\ell^{j+1} > 0$. This is clearly impossible, so we must have $q_{\ell+1}^{j+1} > 0$.

Therefore, $[\mathbf{DC}_k]$ binds and $q_k^{j+1} > 0$ for all $k = 1, \dots, d - 1$. Now consider $k = d$. First, suppose towards a contradiction that $q_d^{j+1} = 0$. Then $[\mathbf{DC}_d]$ and the MLRP hypothesis imply that $q_1^{j+1} \leq \frac{f_{jd}}{f_{j+1,d}} = \min_{1 \leq \ell \leq d} \frac{f_{j\ell}}{f_{j+1,\ell}}$. But then $(q_1^{j+1}, 0, \dots, 0)$ is a solution of $[\mathbf{D}_{j+1,j}]$ such that $q_1^{j+1} \geq \tilde{q}_1^{j+1}$ for all optimal solutions $\tilde{q}^{j+1}$ of $[\mathbf{D}_{j+1,j}]$, delivering the desired contradiction. Thus, $q_d^{j+1} > 0$. Next, suppose towards a contradiction that $[\mathbf{DC}_d]$ is slack. We can then slightly reduce each $q_k^{j+1} > 0$ with $2 \leq k \leq d$ so that all $[\mathbf{DC}_k]$ for $1 \leq k \leq d$ are slack. Thus, there must exist a minimiser of $[\mathbf{D}_{j+1,j}]$ that strictly exceeds q_1^{j+1} , delivering the desired contradiction. Therefore, $[\mathbf{DC}_d]$ binds. $\square$

Lemma I.15. Suppose the type process satisfies PPR and $f_j \neq f_{j+1}$ and $t_{j+1} > 1$. Then there exists an optimal solution q^{j+1} of $[\mathbf{D}_{j+1,j}]$ such that $q_k^{j+1} > 0$ for all $k \leq j + 1$ and $q_k^{j+1} = 0$ for all $k > j + 1$. Moreover, $[\mathbf{DC}_k]$ binds for all $k = 1, \dots, j + 1$, while $[\mathbf{DC}_k]$ is slack for all $k = j + 2, \dots, d$.

Proof. Consider any optimal solution $q^{j+1} = (q_1^{j+1}, \dots, q_d^{j+1})$ of $[\mathbf{D}_{j+1,j}]$ such that $q_1^{j+1} \geq \tilde{q}_1^{j+1}$ for all optimal solutions $\tilde{q}^{j+1}$ of $[\mathbf{D}_{j+1,j}]$.³ Lemma I.12 implies that $q_1^{j+1} \leq 1$. We proceed in steps.

Step 1. We claim that, if $j + 1 < d$, then $(q_1^{j+1}, \dots, q_{j+1}^{j+1}, 0, \dots, 0)$ is also an optimal solution of $[\mathbf{D}_{j+1,j}]$. To see this, note that by PPR and $j + 1 < d$, we have $f_{jd}/f_{j+1,d} = 1$. Since $q_1^{j+1} \leq 1$ by Lemma I.12, it follows that $[\mathbf{DC}_d]$ is still satisfied if we replace $q_d^{j+1} \geq 0$ with $\hat{q}_d^{j+1} = 0$. Since $\hat{q}_d^{j+1} \leq q_d^{j+1}$, this replacement adds (weakly more) slack to $[\mathbf{DC}_k]$ for $k = d - 1$ (and leaves all other constraints unaffected). If $j + 2 = d$, we are done.

³Note that this selection of optimal solutions is only relevant if $v_{j+1} = 0$; in particular, if $v_{j+1} < 0$, then all optimal solutions must specify the same value of q_1^{j+1} .

If $j + 2 < d$, then PPR implies that $f_{j+1,d}/f_{j+2,d} = 1$. Since $q_1^{j+1} \leq 1$ by Lemma I.12, it follows that $[\mathbf{DC}_k]$ for $k \in \{d-1, d\}$ are still satisfied if we replace $(q_{d-1}^{j+1}, q_d^{j+1})$ with $(\hat{q}_{d-1}^{j+1}, \hat{q}_d^{j+1}) = (0, 0)$. Moreover, since $\hat{q}_{d-1}^{j+1} \leq q_{d-1}^{j+1}$, this replacement adds (weakly more) slack to $[\mathbf{DC}_k]$ for $k = d-2$ (and leaves all other constraints unaffected). Proceeding iteratively in this manner, we obtain the claim.

Therefore, if $j + 1 < d$, we henceforth assume that $\mathbf{q}^{j+1}$ is of the form $\mathbf{q}^{j+1} = (q_1^{j+1}, \dots, q_{j+1}^{j+1}, 0, \dots, 0)$, where $q_1^{j+1} \geq \tilde{q}_1^{j+1}$ for all optimal solutions $\tilde{\mathbf{q}}^{j+1}$ of $[\mathbf{D}_{j+1,j}]$.

Step 2. We claim that $q_1^{j+1} < 1$. Suppose, towards a contradiction, that $q_1^{j+1} = 1$.

First, we assert that, if $j \neq 1$, then $q_k^{j+1} = 0$ for all $2 \leq k \leq j$. To see this, note that $j \neq 1$ and PPR imply that $f_{j1}/f_{j+1,1} = 1$. Hence, $\mathbf{q}^{j+1} \geq 0$ and $[\mathbf{DC}_1]$ imply that $q_2^{j+1} = 0$. If $j = 2$, we are done. If $j > 2$, then an analogous argument (using $[\mathbf{DC}_k]$ for $k = 2$) implies that $q_3^{j+1} = 0$ as well. Proceeding inductively, the assertion follows.

Now, Step 1 and the assertion together imply that $\mathbf{q}^{j+1}$ is of the following form: (i) for any $2 \leq k \leq j$, we have $q_k^{j+1} = 0$; (ii) for $k = j+1$, we have $q_{j+1}^{j+1} \geq 0$; and (iii) for any $j+1 < k \leq d$, we have $q_k^{j+1} = 0$. Therefore, since $q_1^{j+1} = 1$ by hypothesis, the constraint $[\mathbf{DC}_k]$ for $k = j$ is equivalent to $f_{j+1,j} \leq f_{jj} - t_{j+1} \cdot q_{j+1}^{j+1}$, and the constraint $[\mathbf{DC}_k]$ for $k = j+1$ is equivalent to $f_{j+1,j+1} \leq f_{j,j+1} + q_{j+1}^{j+1}$. Adding these constraints, we obtain $f_{j+1,j} + f_{j+1,j+1} \leq f_{jj} + f_{j,j+1} + q_{j+1}^{j+1}(1 - t_{j+1})$. Since PPR implies that $f_{j+1,j} + f_{j+1,j+1} = f_{jj} + f_{j,j+1}$, it follows that $q_{j+1}^{j+1}(1 - t_{j+1}) \geq 0$. Since $t_{j+1} > 1$ by hypothesis, this implies $q_{j+1}^{j+1} = 0$. But PPR and the hypothesis that $\mathbf{f}_j \neq \mathbf{f}_{j+1}$ imply that $f_{j,j+1} < f_{j+1,j+1}$, which then violates $[\mathbf{DC}_k]$ for $k = j+1$ (given that $q_{j+1}^{j+1} = 0$).

This yields the desired contradiction. We conclude that $q_1^{j+1} < 1$.

Step 3. We claim that $[\mathbf{DC}_k]$ is slack for all $k \geq j+2$. To see this, note that PPR implies $f_{jk}/f_{j+1,k} = 1$ for all such k , and Step 1 implies that $q_k^{j+1} = 0$ for all such k . Therefore, since $q_1^{j+1} < 1$ by Step 2, the claim follows.

Step 4. We claim that $q_k^{j+1} > 0$ and $[\mathbf{DC}_k]$ binds for all $k = 1, \dots, j+1$.

For the base step, we first claim that $[\mathbf{DC}_1]$ binds. Suppose, towards a contradiction, that this constraint is slack. Then (as in the proof of Lemma I.14), there exists an $\varepsilon > 0$ such that, holding q_1^{j+1} fixed, setting $\hat{q}_2^{j+1} := q_2^{j+1} + \frac{f_{j+1,1} \varepsilon}{t_2}$ and (if $d \geq 3$) inductively setting $\hat{q}_{k+1}^{j+1} := q_{k+1}^{j+1} + \frac{f_{j+1,k}}{t_{k+1}} \left(\frac{q_k^{j+1} - q_{k+1}^{j+1}}{2f_{j+1,k}} \right)$ for all $k = 2, \dots, d$ yields a feasible solution $(q_1^{j+1}, \hat{q}_2^{j+1}, \dots, \hat{q}_d^{j+1})$ at which all of the $[\mathbf{DC}_k]$ constraints are slack. Thus, there must exist a minimiser of $[\mathbf{D}_{j+1,j}]$ that strictly exceeds q_1^{j+1} , delivering the desired contradiction. Therefore, $[\mathbf{DC}_1]$ binds. Now, since $q_1^{j+1} < 1$ (by Step 2) and PPR implies $f_{j1}/f_{j+1,1} \geq 1$ (with equality if $j > 1$), the fact that $[\mathbf{DC}_1]$ binds implies that $q_2^{j+1} > 0$.

Now, since $q_1^{j+1} < 1$ (by Step 2) and PPR implies that $f_{jk}/f_{j+1,k} \geq 1$ for all $1 \leq k \leq j$, an inductive argument analogous to the proof of Lemma I.14 then implies that $q_{k+1}^{j+1} > 0$ and $[\mathbf{DC}_k]$ for all $1 \leq k \leq j$. Suppose, towards a contradiction, that $[\mathbf{DC}_k]$ for $k = j+1$ is slack. Then we can slightly reduce each q_{k+1}^{j+1} for $1 \leq k \leq j$ such that all $[\mathbf{DC}_k]$ for

$1 \leq k \leq j+1$ are slack. Meanwhile, all $[\mathbf{DC}_k]$ for $k \geq j+2$ are slack by Step 3. Thus, there must exist a minimiser of $[\mathbf{D}_{j+1,j}]$ that strictly exceeds q_1^{j+1} , delivering the desired contradiction. Therefore, $[\mathbf{DC}_k]$ for $k = j+1$ also binds. $\square$

Lemma I.16. If $\mathbf{f}_j = \mathbf{f}_{j+1}$, there exists an optimal solution $\mathbf{q}^{j+1}$ of $[\mathbf{D}_{j+1,j}]$ with $q_1^{j+1} = 1$. Moreover, $[\mathbf{DC}_k]$ binds for all $k = 1, \dots, d$.

Proof. Since $\mathbf{f}_j = \mathbf{f}_{j+1}$ implies that $f_{jk}/f_{j+1,k} = 1$ for all $1 \leq k \leq d$, $\hat{\mathbf{q}}^{j+1} = (1, 0, \dots, 0)$ is a feasible solution of $[\mathbf{D}_{j+1,j}]$ in which all $[\mathbf{DC}_k]$ constraints bind. Since Lemma I.12 implies that $q_1^{j+1} \leq 1$ in every feasible solution and we have $\hat{q}_1^{j+1} = 1$ by construction, $\hat{\mathbf{q}}^{j+1}$ is an optimal solution of $[\mathbf{D}_{j+1,j}]$. $\square$

Lemma I.17. Fix $\mathbf{t} \in \mathbb{R}_+^{d-1}$ such that $\mathbf{t} \geq \mathbf{1}$ and $t_2 \cdots t_d > 1$, and let $K_{\mathbf{t}}$ be defined as in [I.6] above. If the type process satisfies MLRP, $\mathcal{B}^\dagger(K_{\mathbf{t}}) \neq K_{\mathbf{t}}$.

Proof. For each dual problem $[\mathbf{D}_{j+1,j}]$, by Lemmas I.14 and I.16, the dual constraints $[\mathbf{DC}_k]$ bind for all $k = 1, \dots, d$. A little algebra (eliminating q_k^{j+1} for $k = 2, \dots, d$ by multiplying each side of $[\mathbf{DC}_k]$ by $f_{j+1,k} \prod_{i=2}^d t_i$ for $k = 2, \dots, d$, multiply both sides of $[\mathbf{DC}_1]$ by $f_{j+1,1}$, and then adding) tells us that

$$q_1^{j+1} = \frac{f_{j1} + t_2 f_{j2} + \cdots + t_2 \cdots t_d f_{jd}}{f_{j+1,1} + t_2 f_{j+1,2} + \cdots + t_2 \cdots t_d f_{j+1,d}}$$

for all $j = 1, \dots, d-1$. Recall that the maximal value for $[\mathbf{P}_{j+1,j}]$ is v_j . By the duality theorem for Linear Programming, it must be that $v_j = q_1^{j+1} v_{j+1}$, where the latter is the value of $[\mathbf{D}_{j+1,j}]$. In other words, any $\mathbf{v}$ with $v_{j+1} \geq t'_{j+1} v_j$ is implementable, where $t'_{j+1} = 1/q_1^{j+1}$. This implies $\mathcal{B}^\dagger(K_{\mathbf{t}}) = K_{\mathbf{t}'}$, where $\mathbf{t}' = (t'_2, \dots, t'_d)$. It is easy to see that

$$t'_2(\mathbf{t}) \cdot t'_3(\mathbf{t}) \cdots t'_d(\mathbf{t}) = \frac{f_{d1} + t_2 f_{d2} + \cdots + t_2 \cdots t_d f_{dd}}{f_{11} + t_2 f_{12} + \cdots + t_2 \cdots t_d f_{1d}}$$

To show that $\mathcal{B}^\dagger(K_{\mathbf{t}}) \neq K_{\mathbf{t}}$, it suffices to show that $t'_2(\mathbf{t}) \cdot t'_3(\mathbf{t}) \cdots t'_d(\mathbf{t}) \neq t_2 \cdots t_d$. But this follows immediately from the observation that

$$\frac{1}{t_2 \cdots t_d} f_{d1} + \frac{1}{t_3 \cdots t_d} f_{d2} + \cdots + f_{dd} < 1 < f_{11} + t_2 f_{12} + \cdots + t_2 \cdots t_d f_{1d}$$

which is true because $\mathbf{t} \geq \mathbf{1}$ and $t_2 \cdots t_d > 1$. $\square$

Lemma I.18. Fix $\mathbf{t} \in \mathbb{R}_+^{d-1}$ such that $\mathbf{t} \geq \mathbf{1}$ and $t_2 \cdots t_d > 1$, and let $K_{\mathbf{t}}$ be defined as in [I.6] above. If the type process satisfies PPR, $\mathcal{B}^\dagger(K_{\mathbf{t}}) \neq K_{\mathbf{t}}$.

Proof. For each dual problem $[\mathbf{D}_{j+1,j}]$, by Lemmas I.15 and I.16, the dual constraints $[\mathbf{DC}_k]$

bind for all $k = 1, \dots, j + 1$. A little algebra tells us that

$$q_1^{j+1} = \frac{f_{j1} + t_2 f_{j2} + \dots + t_2 \dots t_{j+1} f_{j,j+1}}{f_{j+1,1} + t_2 f_{j+1,2} + \dots + t_2 \dots t_{j+1} f_{j+1,j+1}}$$

Recall that the maximal value for $[\mathbf{P}_{j+1,j}]$ is v_j . By the duality theorem for Linear Programming, it must be that $v_j = q_1^{j+1} v_{j+1}$, where the latter is the value of $[\mathbf{D}_{j+1,j}]$. In other words, any $\mathbf{v}$ with $v_{j+1} \geq t'_{j+1} v_j$ is implementable, where $t'_{j+1} = 1/q_1^{j+1}$. This implies $\mathcal{B}^\dagger(K_{\mathbf{t}}) = K_{\mathbf{t}'}$, where $\mathbf{t}' = (t'_2, \dots, t'_d)$. To show that $\mathcal{B}^\dagger(K_{\mathbf{t}}) \neq K_{\mathbf{t}}$, it suffices to show that if $t_{j+1} > 1$, then $t'_{j+1} < t_{j+1}$.

To see that this is the case, observe that $t'_{j+1} < t_{j+1}$ if, and only if, $\varphi(t_{j+1}) > 0$, where

$$\begin{aligned} \varphi(t_{j+1}) := & t_{j+1}^2 [t_2 \dots t_j f_{j,j+1}] + t_{j+1} [f_{j1} + t_2 f_{j2} + \dots + t_2 \dots t_j f_{jj} - t_2 \dots t_j f_{j+1,j+1}] \\ & - [f_{j+1,1} + t_2 f_{j+1,2} + \dots + t_2 \dots t_j f_{j+1,j}] \end{aligned}$$

The function φ is a quadratic in t_{j+1} . It is clear that $\varphi(0) < 0$, and a simple calculation establishes that $\varphi(1) = 1$. Moreover, φ is convex, which implies $\varphi(t_{j+1}) > 0$ whenever $t_{j+1} > 1$, which establishes the claim. $\square$

I.2.3. Iterates of $\mathcal{B}$ in the CARA Case

Define inductively the sequence of sets $K^{(n)}$ as follows: $K^{(0)} := \mathbb{R}_-^d$, and for all $n \geq 1$, $K^{(n)} := \mathcal{B}(K^{(n-1)})$.

Lemma I.19. The sequence $(K^{(n)})$ satisfies:

- (a) For each $n \in \mathbb{N}$, $K^{(n)}$ is a closed, convex cone in $\mathbb{R}_-^d$.
- (b) $K^{(n)} \subseteq K^{(n-1)}$ for all $n \geq 1$.
- (c) The largest fixed point of $\mathcal{B}$, $K_* = \lim_{n \rightarrow \infty} K^{(n)}$ and is a closed, convex cone.

Proof. Part (a) follows from induction and Lemma I.4. Part (b) follows from monotonicity of $\mathcal{B}$. The limit statement in part (c) follows from the simple observation that $\mathcal{B}$ is order-continuous on the complete lattice $(2^{\mathbb{R}_-^d}, \supseteq)$ and arguments in Section 4 of Echenique (2005). That K_* is a closed, convex cone follows from the limit statement and monotonicity of the sequence as described in part (b). $\square$

We would like to establish analogues of Lemma I.12 and Corollary I.13 for this more general setting. To do so, we show how to compute iterates of $\mathcal{B}$ through solutions of conic programs. (As will become clear, due to the multiple incentive constraints, in contrast to the LP analysis in the previous section, it is no longer clear that the iterates of $\mathcal{B}$ are *polyhedral* cones; even if they are, the generating polyhedra are more difficult to explicitly describe. Since we will not need to compute the iterates in closed form, we adopt the more general approach of conic, instead of linear, programming.)

Given a convex cone $C \subseteq \mathbb{R}^n$, the *dual cone* is $C^* := \{y \in \mathbb{R}^n : y \cdot x \geq 0 \ \forall x \in C\}$. We will use the following version of Conic Programming Strong Duality in the sequel.⁴

Theorem 5 (Conic Programming Strong Duality). *Let $J \subseteq \mathbb{R}^n$ and $L \subseteq \mathbb{R}^m$ be closed convex cones, and let J^* and L^* denote the respective dual cones. Let $b \in \mathbb{R}^m$ and $A \in \mathbb{R}^{m \times n}$. Consider the primal conic program*

$$[\mathbf{P}^{\text{cone}}] \quad \max_y \quad c \cdot y \quad \text{subject to } b - Ay \in L \text{ and } y \in J$$

and the dual conic program

$$[\mathbf{D}^{\text{cone}}] \quad \min_q \quad b \cdot q \quad \text{subject to } A^\top q - c \in J^* \text{ and } q \in L^*$$

Suppose that $[\mathbf{P}^{\text{cone}}]$ satisfies the Slater condition, i.e., there exists $z \in J$ such that $b - Az \in \text{int}L$. Then, if $[\mathbf{P}^{\text{cone}}]$ is feasible and has finite value:

- (a) The dual problem $[\mathbf{D}^{\text{cone}}]$ is feasible.
- (b) The values of $[\mathbf{P}^{\text{cone}}]$ and $[\mathbf{D}^{\text{cone}}]$ are equal (and finite).

We will also use the following piece of notation in the sequel: for any $v \in \mathbb{R}_-^d$ and $j = 1, \dots, d-1$, define $v_{>j} := (v_{j+1}, \dots, v_d) \in \mathbb{R}_-^{d-j}$.

Given some closed, convex cone $K \subseteq \mathbb{R}_+^d$, some $j = 1, \dots, d-1$ and $v_{>j} \in \mathbb{R}_+^{d-j}$, consider the following conic programming problem:

$$\begin{aligned}
[\mathbf{P}_j^{\text{full}}] \quad & \max_{u_j, \mathbf{w}_j} \quad u_j + \alpha[f_{j1}w_{j1} + f_{j2}w_{j2} + \dots + f_{jd}w_{jd}] \\
\text{s.t.} \quad & \frac{\theta_{j+1}}{\theta_j} \cdot u_j + \alpha[f_{j+1,1}w_{j,1} + \dots + f_{j+1,d}w_{j,d}] \leq v_{j+1} \\
& \vdots \\
& \frac{\theta_d}{\theta_j} \cdot u_j + \alpha[f_{d,1}w_{j,1} + \dots + f_{d,d}w_{j,d}] \leq v_d \\
& u_j \leq 0 \\
& \mathbf{w}_j \in K
\end{aligned}$$

It is easy to see that this is a primal conic programming problem of the form described in $[\mathbf{P}^{\text{cone}}]$, with $y = (u_1, \mathbf{w}_1)$, $b = v_{>j}$, $L = \mathbb{R}_+^d$, $J = \mathbb{R}_- \times K$, and the vector c and matrix A easy to read off from the problem. Thus, the dual cones are $L^* = L$ and $J^* = \mathbb{R}_- \times K^*$. It

⁴This statement of conic duality is from Theorem 4.7.1 of Gärtner and Matoušek (2012, p. 62).

follows that the dual conic program may be written as

$$\begin{aligned}
[\mathbf{D}_j^{\text{full}}] \quad & \min_{\mathbf{q}^j \in \mathbb{R}^{d-j}} \sum_{k=j+1}^d q_{k-j}^j v_k \\
[\mathbf{DC}_0^{\text{full}}] \quad & \text{s.t.} \quad \sum_{k=j+1}^d q_{k-j}^j \theta_k \leq \theta_j \\
[\mathbf{DC}_{K^*}^{\text{full}}] \quad & \alpha [\mathbf{f}_{j+1}, \mathbf{f}_{j+2}, \dots, \mathbf{f}_d] \mathbf{q}^j - \alpha \mathbf{f}_j \in K^* \\
& \mathbf{q}^j \geq 0
\end{aligned}$$

where $[\mathbf{f}_{j+1}, \mathbf{f}_{j+2}, \dots, \mathbf{f}_d]$ denotes the $d \times (d-j)$ matrix whose k th column is given by the (column) vector of transition probabilities from state $j+k$, i.e., $\mathbf{f}_{j+k} = (f_{j+k,1}, \dots, f_{j+k,d})^\top$.

Let $MP_j : \mathbb{R}_i^{d-j} \times \mathcal{K} \rightrightarrows \mathbb{R}_-^{d+1}$ denote the argmax correspondence from $[\mathbf{P}_j^{\text{full}}]$, as a function of $(\mathbf{v}_{>j}, K)$. (Recall that $\mathcal{K}$ denotes the space of closed convex cones in $\mathbb{R}_-^d$.) Let $VP_j : \mathbb{R}_i^{d-j} \times \mathcal{K} \rightarrow \mathbb{R}_-$ denote the value function for $[\mathbf{P}_j^{\text{full}}]$. Similarly, let $MD_j : \mathbb{R}_i^{d-j} \times \mathcal{K} \rightrightarrows \mathbb{R}_+^{d-j}$ and $VD_j : \mathbb{R}_i^{d-j} \times \mathcal{K} \rightarrow \mathbb{R}_-$ denote the argmin correspondence and value function from $[\mathbf{D}_j^{\text{full}}]$.

Lemma I.20. Let $j \leq d-1$ be given. For any $\{0\} \neq K \in \mathcal{K}$, the following hold:

- (a) For any $\mathbf{v}_{>j} \in \mathbb{R}_i^{d-j}$, the Slater condition and strong conic programming duality (conclusions (a) and (b) of Theorem 5) hold.
- (b) $VP_j(\cdot, K)$ and $VD_j(\cdot, K)$ are well-defined (i.e., finite-valued) continuous, concave, and HD1.
- (c) $MP_j(\cdot, K)$ and $MD_j(\cdot, K)$ are nonempty-valued and upper hemi-continuous. $MD_j(\cdot, K)$ is HD0.
- (d) $VP_j(\mathbf{v}_{>j}, K) = VD_j(\mathbf{v}_{>j}, K) < 0$ if and only if $\mathbf{v}_{>j} \neq 0$.

Proof. First, observe that $[\mathbf{P}_j^{\text{full}}]$ is clearly feasible and satisfies the Slater condition, as it is always possible to set $u_j = 0$ and scale up $\mathbf{w}_j$ along some ray contained in K so that all of the incentive constraints hold as strict inequalities. Note that the following choice is always feasible in $[\mathbf{P}_j^{\text{full}}]$: $\mathbf{w}_j = \mathbf{0}$ and $u_j = \theta_j \cdot \min_{k=j+1, \dots, d} [v_k / \theta_k]$. This policy has finite value $\theta_j \cdot \min_{k=j+1, \dots, d} [v_k / \theta_k]$. (The value is always bounded above by 0.) Thus, it is without loss of optimality to restrict attention to $(u_j, \mathbf{w}_j) \in R(\mathbf{v}_{>j})$ where

$$R(\mathbf{v}_{>j}) := \left\{ (u_j, \mathbf{w}_j) \in \mathbb{R}_-^{d+1} : u_j + \alpha [f_{j1}w_{j1} + f_{j2}w_{j2} + \dots + f_{jd}w_{jd}] \geq \theta_j \cdot \min_{k=j+1, \dots, d} [v_k / \theta_k] \right\}$$

is nonempty and compact. Define the correspondence $M : \mathbb{R}_i^{d-j} \times \mathcal{K} \rightrightarrows \mathbb{R}_-^{d+1}$ by

$$M(\mathbf{v}_{>j}, K) := \left\{ (u_j, \mathbf{w}_j) \in R(\mathbf{v}_{>j}) : (u_j, \mathbf{w}_j) \text{ is feasible in } [\mathbf{P}_j^{\text{full}}] \text{ at } (\mathbf{v}_{>j}, K) \right\}$$

It is easy to see that, for fixed K , $M(\cdot, K)$ is nonempty- and compact-valued and continuous.

Part (a) thus follows from Theorem 5. The constraint set in $[\mathbf{D}_j^{\text{full}}]$ independent of $\mathbf{v}_{>j}$ and is clearly compact (due to $[\mathbf{DC}_0^{\text{full}}]$ and $\mathbf{q}^j \geq 0$). Part (c) thus follows from these observations and Berge's Theorem of the Maximum, as does continuity of the value functions stated in part (b). Concavity follows from the fact that $VD_j(\cdot, K)$ is the lower envelope of linear functions and strong duality from part (a). HD1 follows from the simple observation that $MD_j(\cdot, K)$ must be HD0, i.e., $\mathbf{q}^j \in MD_j(\mathbf{v}_{>j}, K)$ if and only if $\mathbf{q}^j \in MD_j(a\mathbf{v}_{>j}, K)$ for all $a \geq 0$.

Finally, part (d) is obvious upon inspection of the constraints in $[\mathbf{P}_j^{\text{full}}]$. $\square$

Now we may state the analogue of Lemma I.12.

Lemma I.21. The sequence $(K^{(n)})$ satisfies:

$$[\text{I.7}] \quad K^{(n)} = \{ \mathbf{v} \in \mathbb{R}_-^d : v_j \leq VP_j(\mathbf{v}_{>j}, K^{(n-1)}) \text{ for all } j = 1, \dots, d-1 \}$$

Thus, the largest fixed point of $\mathcal{B}$, K_* , satisfies

$$[\text{I.8}] \quad K_* = \{ \mathbf{v} \in \mathbb{R}_-^d : v_j \leq VP_j(\mathbf{v}_{>j}, K_*) \text{ for all } j = 1, \dots, d-1 \}$$

Proof. Immediate from the definitions. $\square$

We remark that parts (b) and (c) of Lemma I.20 imply that $K^{(n)}$ and K_* indeed satisfy all the properties laid out in Lemma I.19.

Define the following sets:

$$[\text{I.9}] \quad A_j := \{ \mathbf{v} \in \mathbb{R}_-^d : v_j = 0 \} \quad A := \bigcup_{j=1}^d A_j$$

$$[\text{I.10}] \quad B_j := \{ \mathbf{v} \in \mathbb{R}_-^d : v_j = VP_j(\mathbf{v}_{>j}, K_*) \} \quad B := \bigcup_{j=1}^{d-1} B_j$$

Lemma I.22. The boundary of K_* satisfies $\text{bd } K_* \subseteq A \cup B$. The interior of K_* satisfies

$$[\text{I.11}] \quad \text{int } K_* = \{ \mathbf{v} \in \mathbb{R}_-^d : v_j < VP_j(\mathbf{v}_{>j}, K_*) \text{ for all } j = 1, \dots, d-1 \}$$

Proof. Immediate from Lemma I.21 and the fact that, by part (b) of Lemma I.20, each $VP_j(\cdot, K_*)$ is continuous. $\square$

Finally, the desired analogue to Corollary I.13.

Lemma I.23. For any $K \in \{K^{(n)}\}_{n \in \mathbb{N}} \cup \{K_*\}$ and $(u_j, \mathbf{w}_j) \in MP_j(\mathbf{v}_{>j}, K)$, we have $u_j = 0$.

Proof. Let any such K be given. By Proposition I.8 and Lemma I.19, K is a closed convex cone satisfying $\text{cl } V_d \subseteq K \subseteq \mathbb{R}_-^d$. Suppose, towards a contradiction, that $(u_j, \mathbf{w}_j) \in MP_j(\mathbf{v}_{>j}, K)$ and yet $u_j < 0$. Consider $(u'_j, \mathbf{w}'_j)$ defined as $u'_j := u_j + \alpha\varepsilon$ and $\mathbf{w}'_j := \mathbf{w}_j - \varepsilon\mathbf{1}$ for some $\varepsilon > 0$ sufficiently small that $u'_j < 0$.

First, we claim that $(u'_j, \mathbf{w}'_j)$ is feasible in $[\mathbf{P}_j^{\text{full}}]$. To see this, note that $-\varepsilon\mathbf{1} \in \text{cl } V_d \subseteq K$ and $\mathbf{w}_j \in K$ by construction, and that K is closed under addition because it is a convex cone. Thus, $\mathbf{w}'_j \in K$. Now, for each $k \geq j+1$, the k -th IC constraint in $[\mathbf{P}_j^{\text{full}}]$ reads

$$\begin{aligned} \frac{\theta_k}{\theta_j} u'_j + \alpha \mathbf{E}^{\mathbf{f}_k}[\mathbf{w}'_j] &= \frac{\theta_k}{\theta_j} (u_j + \alpha\varepsilon) + \alpha \mathbf{E}^{\mathbf{f}_k}[\mathbf{w}_j - \varepsilon\mathbf{1}] \\ &\leq v_k - \alpha\varepsilon(1 - \theta_k/\theta_j) \\ &< v_k \end{aligned}$$

where the first line is by definition, the second line holds because $(u_j, \mathbf{w}_j)$ is feasible by assumption, and the final line holds because $\theta_k < \theta_j$. Thus, $(u'_j, \mathbf{w}'_j)$ is feasible in $[\mathbf{P}_j^{\text{full}}]$.

Next, note that $(u'_j, \mathbf{w}'_j)$ attains the same value as $(u_j, \mathbf{w}_j)$, because $u'_j + \alpha \mathbf{E}^{\mathbf{f}_j}[\mathbf{w}'_j] = u_j + \alpha\varepsilon + \alpha \mathbf{E}^{\mathbf{f}_j}[\mathbf{w}_j - \varepsilon\mathbf{1}] = u_j + \alpha \mathbf{E}^{\mathbf{f}_j}[\mathbf{w}_j]$ by construction. Thus, $(u'_j, \mathbf{w}'_j) \in MP_j(\mathbf{v}_{>j}, K)$.

But, then, since $u'_j < 0$ and all of the IC constraints are slack under $(u'_j, \mathbf{w}'_j)$ by construction (as show above), there is some $\delta > 0$ such that $(u'_j + \delta, \mathbf{w}'_j)$ is feasible in $[\mathbf{P}_j^{\text{full}}]$ and attains strictly higher value. This is the desired contradiction. It follows that $u_j = 0$, completing the proof. $\square$

I.2.4. Iterates of $\mathcal{B}$ in the General Case

We now turn to the case of general utility functions $U(\cdot)$ satisfying Assumption DARA. Unlike the case of CARA utility, it is no longer immediate that $\mathcal{B}$ maps convex cones to convex cones. Below, we show that this is nevertheless true.

Given some closed, convex cone $K \subseteq \mathbb{R}_-^d$, some $j = 1, \dots, d-1$, and $\mathbf{v}_{>j} \in \mathbb{R}_-^{d-j}$, consider the following optimization problem:

$$\begin{aligned} [\mathbf{P}_j^{\text{DARA}}] \quad & \max_{u_j, \mathbf{w}_j} \quad u_j + \alpha[f_{j1}w_{j1} + f_{j2}w_{j2} + \dots + f_{jd}w_{jd}] \\ \text{s.t.} \quad & \psi(u_j, j+1, j) + \alpha[f_{j+1,1}w_{j,1} + \dots + f_{j+1,d}w_{j,d}] \leq v_{j+1} \\ & \vdots \\ & \psi(u_j, d, j) + \alpha[f_{d,1}w_{j,1} + \dots + f_{d,d}w_{j,d}] \leq v_d \\ & u_j \leq 0 \\ & \mathbf{w}_j \in K \end{aligned}$$

Note that problem $[\mathbf{P}_j^{\text{DARA}}]$ is a strict generalization of $[\mathbf{P}_j^{\text{full}}]$. It is also a *concave* optimization problem because, by Lemma I.1, the constraint set is closed and convex.

Proposition I.24. Let $K \in \{K^{(n)}\}_{n \in \mathbb{N}} \cup \{K_*\}$. Then, the set $\mathcal{B}(K^{(n)})$ is independent of the discount factor $\alpha \in (0, 1)$ and the utility function $U(\cdot)$ (which satisfies Assumption 1).

Proof. To begin, let $n = 0$ so that $K = K^{(0)} = \mathbb{R}_-^d$. Clearly, K is independent of α and U .

Suppose, towards a contradiction, that $(u_j, \mathbf{w}_j)$ is an optimal solution to $[\mathbf{p}_j^{\text{DARA}}]$ and $u_j < 0$. Consider $(u'_j, \mathbf{w}'_j)$ defined as $u'_j := u_j + \alpha\varepsilon$ and $\mathbf{w}'_j := \mathbf{w}_j - \varepsilon\mathbf{1}$ for some $\varepsilon > 0$ sufficiently small that $u'_j < 0$. We proceed analogously to the proof of Lemma I.23.

First, we claim that $(u'_j, \mathbf{w}'_j)$ is feasible in $[\mathbf{p}_j^{\text{DARA}}]$. To see this, note that $-\varepsilon\mathbf{1} \in \text{cl } V_d \subseteq K$ and $\mathbf{w}_j \in K$ by construction, and that K is closed under addition because it is a convex cone. Thus, $\mathbf{w}'_j \in K$. Now, for each $k \geq j + 1$, the k -th IC constraint in $[\mathbf{p}_j^{\text{DARA}}]$ reads

$$\begin{aligned} \psi(u'_j, k, j) + \alpha \mathbf{E}^{\mathbf{f}_k}[\mathbf{w}'_j] &\leq \psi(u_j, k, j) + \alpha\varepsilon \cdot \psi'(u'_j, k, j) + \alpha \mathbf{E}^{\mathbf{f}_k}[\mathbf{w}_j - \varepsilon\mathbf{1}] \\ &\leq v_k - \alpha\varepsilon(1 - \psi'(u'_j, k, j)) \\ &< v_k \end{aligned}$$

where the first line is by definition of $(u'_j, \mathbf{w}'_j)$ and the convexity of $\psi(\cdot, k, j)$ (Lemma I.1), the second line holds because $(u_j, \mathbf{w}_j)$ is feasible by assumption, and the final line holds because $\psi'(u'_j, k, j) < 1$ (Lemma I.1). Thus, $(u'_j, \mathbf{w}'_j)$ is feasible in $[\mathbf{p}_j^{\text{DARA}}]$.

Next, note that $(u'_j, \mathbf{w}'_j)$ attains the same value as $(u_j, \mathbf{w}_j)$, because $u'_j + \alpha \mathbf{E}^{\mathbf{f}_j}[\mathbf{w}'_j] = u_j + \alpha\varepsilon + \alpha \mathbf{E}^{\mathbf{f}_j}[\mathbf{w}_j - \varepsilon\mathbf{1}] = u_j + \alpha \mathbf{E}^{\mathbf{f}_j}[\mathbf{w}_j]$ by construction. Thus, $(u'_j, \mathbf{w}'_j)$ is also an optimal solution of $[\mathbf{p}_j^{\text{DARA}}]$. But, then, since $u'_j < 0$ and all of the IC constraints are slack under $(u'_j, \mathbf{w}'_j)$ by construction (as show above), there is some $\delta > 0$ such that $(u'_j + \delta, \mathbf{w}'_j)$ is feasible in $[\mathbf{p}_j^{\text{DARA}}]$ and attains strictly higher value. This is the desired contradiction.

Therefore, it holds that $u_j = 0$ in any optimal solution of $[\mathbf{p}_j^{\text{DARA}}]$ for $K = K^{(0)} = \mathbb{R}_-^d$. By Lemma I.23, the same is true of every optimal solution of $[\mathbf{p}_j^{\text{full}}]$, given any CARA utility function. Moreover, since $\psi(0, k, j) := \lim_{u \rightarrow 0} \psi(u, k, j) = 0$ for all $k \geq j + 1$ (Lemma I.1), the constraints and objectives in $[\mathbf{p}_j^{\text{DARA}}]$ and $[\mathbf{p}_j^{\text{full}}]$ are identical whenever u_j ; in particular, they depend only on the transition probabilities $\{\mathbf{f}_k\}_{k=j+1}^d$ and the discount factor α , but not on the utility function U . It follows that the optimal solutions of $[\mathbf{p}_j^{\text{DARA}}]$ and $[\mathbf{p}_j^{\text{full}}]$ coincide, and are independent of U , at $K = K^{(0)} = \mathbb{R}_-^d$.

Proceeding inductively in $n \in \mathbb{N}$, the above argument, Proposition I.8, and Lemma I.19 together imply that each $K^{(n)}$ is a closed convex cone satisfying $\text{cl } V_d \subseteq K^{(n)} \subseteq \mathbb{R}_-^d$, and such that $K^{(n)}$ does not depend on U . Since Proposition I.8 and Lemma I.19 then imply that $K_* = \lim_{n \rightarrow \infty} K^{(n)}$ is also closed convex cone satisfying $\text{cl } V_d \subseteq K^{(n)} \subseteq \mathbb{R}_-^d$, the same argument also yields that $\mathcal{B}(K_*) = K_*$ is the independent of U .

Finally, we show independence with respect to the discount factor $\alpha \in (0, 1)$. By the above, each $K \in \{K^{(n)}\}_{n=0}^\infty \cup \{K_*\}$ is a convex cone; hence, the objective and constraints in $[\mathbf{p}_j^{\text{DARA}}]$ are HD1 in $(\mathbf{v}, \mathbf{w}_j)$ when $u_j = 0$. Moreover, by the above, the optimal solutions of each $[\mathbf{p}_j^{\text{DARA}}]$ involves $u_j = 0$. Since $K^{(0)} = \mathbb{R}_-^d$ is clearly independent of α , the result

then follows from induction in n and taking the $n \rightarrow \infty$ limit. $\square$

Proposition I.25. The following hold:

- (a) K_* , the largest fixed point of $\mathcal{B}$, is independent of α and U .
- (b) $D = \mathcal{B}_\circ(K_*) = \text{int}K_*$. Thus, D is an open, convex cone and is independent of α and U .

Proof. Part (a) follows from Proposition I.24, Lemma I.21, and part (c) of Lemma I.19.

For part (b), consider problem $[\mathbf{p}_j^{\text{DARA}}]$ given K_* . By Proposition I.24, K_* is characterized by Lemmas I.21 and I.22. Recall the sets A and B defined in [I.9] and [I.10] before Lemma I.22. Clearly, $\mathbf{v} \in A$ implies that $\mathbf{v} \notin \mathcal{B}_\circ(K_*)$. Consider $\mathbf{v} \in B$. It follows from the proof of Proposition I.24 that any menu $(u_j, \mathbf{w}_j)_{j \in S} \in (\mathbb{R}_- \times K_*)^d$ that implements $\mathbf{v}$ must satisfy $u_j = 0$ for some $j = 1, \dots, d-1$. Thus, $\mathbf{v} \notin \mathcal{B}_\circ(K_*)$.

It therefore follows from Lemma I.22 that $\mathcal{B}_\circ(K_*) \subseteq \text{int}K_*$. We claim that $\mathcal{B}_\circ(\text{int}K_*) \supseteq \text{int}K_*$. Let $\mathbf{v} \in K_*$, so that by Lemma I.22 $v_d < 0$ and $v_j < VP_j(\mathbf{v}_{>j}, K_*)$ for all $j = 1, \dots, d-1$. Consider the concave program $[\mathbf{p}_j^{\text{DARA}}]$ (given this $\mathbf{v}$ and K_*). The constraint set is a convex set with nonempty interior (this follows from the same simple arguments used to establish the Slater condition in the proofs of part (a) of Lemma I.20 and Proposition I.24), and thus for any feasible $(u_j, \mathbf{w}_j)$ there exists a sequence $(u_j^n, \mathbf{w}_j^n) \rightarrow (u_j, \mathbf{w}_j)$ such that each $(u_j^n, \mathbf{w}_j^n)$ lies in the interior of the feasible set. The objective function is continuous, so there exists some $N \in \mathbb{N}$ such that for any $n \geq N$ we have $v_j < u_j^n + \alpha \mathbf{E}^{f_j}[\mathbf{w}_j^n]$ for all $j = 1, \dots, d-1$. Thus, $\mathbf{v} \in \mathcal{B}_\circ(\text{int}K_*)$. Thus, we have proved the claim.

Now, monotonicity of $\mathcal{B}_\circ$ implies that $\mathcal{B}_\circ(K_*) = \mathcal{B}_\circ(\text{int}K_*) = \text{int}K_*$ and thus also that $D = \text{int}K_*$, which completes the proof. $\square$

I.3. A Useful Suboptimal Contract

In this section, we consider the case in which the agent has CARA utility and the type process satisfies either MLRP or PPR. The goal is to construct a (suboptimal) recursive contract $\zeta \in \Xi$ that is [TVC]-implementable starting from any $\mathbf{v} \in D$.⁵ This serves as an intermediate step towards proving both part (f) of Theorem 3 (see Section I.4 below). We also show that ζ generates well-defined and finite value for the principal starting from any $\mathbf{v} \in D$, which is needed to prove part (b) of Lemma B.2 in Appendix B.2; this is nontrivial because the principal's flow cost function $C(\cdot)$ is unbounded both above and below.

⁵Note that, with a slight abuse of notation, here we use $\zeta \in \Xi$ to denote a specific contract that is feasible in the second-best problem (with IC constraints), rather than a generic element of Ξ^{FB} , the set of full-information feasible contracts (cf. Section 3.1 and SA-F of BKL).

For simplicity, for the remainder of (only) the present section, we restrict attention to the $d = 2$ case; extending to the general case for $d \geq 3$ is straightforward but notationally very cumbersome. Let $\beta := f_{22} - f_{12} = f_{11} - f_{21}$. (The equality $f_{22} - f_{12} = f_{11} - f_{21}$ is an identity that follows from $f_{11} + f_{12} = f_{21} + f_{22} = 1$.) Because we are considering the MLRP/PPR case types and the process is fully connected by Assumption [Markov](#), we have $\beta \in [0, 1)$. Because we are considering CARA utility of the form $U(c) := -e^{-\rho c}$ for some $\rho > 0$, it is convenient to work directly with the variables $x_i := u_i/\theta_i$ where $\theta_i := e^{-\rho \omega_i}$. Thus, with a slight abuse of notation, we will let recursive contracts specify flow allocations in terms of the x_i instead of the u_i .

I.3.1. Some Derived Parameters and Relations Between Them

Let $t_0 := \theta_2/\theta_1$. It is easy to see that $t_0 \in (0, 1)$. Define $\eta_1 := f_{21} - t_0 f_{11}$ and $\eta_2 := f_{22} - t_0 f_{12}$. It is easy to see that $\eta_2 > 0$. However, we do not have sufficient information to sign η_1 . Let $\lambda_* := \alpha f_{21}/(1 - \alpha f_{22})$ and $\lambda^* := (\alpha \eta_1 + t_0)/(1 - \alpha \eta_2)$.

We shall now record some facts about these newly defined parameters.

Lemma I.26. We have $\eta_1 + t_0 \eta_2 > 0$. In particular, the quadratic $\psi(t) = -f_{12}t^2 + (f_{22} - f_{11})t + f_{21} > 0$ for all $t \in (0, 1)$, and $\psi(t_0) = \eta_1 + t_0 \eta_2$.

Proof. Of course, if $\eta_1 \geq 0$, then the inequality always holds. We shall show that it holds even if $\eta_1 < 0$.

Notice that

$$\begin{aligned} \eta_1 + t_0 \eta_2 &= f_{21} - t_0 f_{11} + t_0(f_{22} - t_0 f_{12}) \\ &= -f_{12}t_0^2 + (f_{22} - f_{11})t_0 + f_{21} \\ &=: \psi(t_0) \end{aligned}$$

where $\psi(t)$ is the quadratic polynomial defined above. Notice that $\psi(1) = 0$, while $\psi(0) = f_{21} > 0$. Also observe that $\psi'(t) = -2f_{12}t + (f_{22} - f_{11})$, which implies that $\psi'(1) = -2f_{12} + f_{22} - f_{11} = \beta - 1 < 0$.

Because ψ is a quadratic, it can have at most two real zeros. But 1 is a zero of ψ , which implies that the other zero is real and (strictly) negative. Therefore, for all $t \in [0, 1)$, $\psi(t) > 0$, which completes the proof. $\square$

Lemma I.27. The following inequalities hold:

- (a) $0 < \lambda_*, \lambda^* < 1$.
- (b) $t_0 < \lambda^*$.
- (c) $\lambda_* < \lambda^*$.

Proof. Part (a) follows immediately from basic assumptions in the model, so we shall only establish parts (b) and (c).

To see (b), notice that $t_0 \leq \lambda^*$ if, and only if, $t_0 - t_0\alpha\eta_2 \leq \alpha\eta_1 + t_0$, which holds if, and only if, $\alpha(\eta_1 + t_0\eta_2) > 0$, which is true by Lemma I.26.

To see (c), notice that $\alpha f_{21}/(1 - \alpha f_{22}) < (\alpha\eta_1 + t_0)/(1 - \alpha\eta_2)$ if, and only if, $\alpha f_{21} - \alpha^2 f_{21}\eta_2 < \alpha\eta_1 + t_0 - \alpha^2 f_{22}\eta_1 - \alpha t_0 f_{22}$. This is equivalent to requiring that

$$\begin{aligned}\xi(\alpha) &:= \alpha^2(f_{21}\eta_2 - f_{22}\eta_1) + \alpha(\eta_1 - f_{21} + f_{22}t_0) + t_0 \\ &= \alpha^2 t_0(f_{22}f_{11} - f_{12}f_{21}) + \alpha t_0(f_{22} - f_{11}) + t_0 > 0\end{aligned}$$

It is easy to see that $f_{22}f_{11} \geq f_{12}f_{21}$. We also have $\alpha t_0(f_{22} - f_{11}) + t_0 = t_0[1 + \alpha(f_{22} - f_{11})] > 0$ because $1 + \alpha(f_{22} - f_{11}) > 0$. This implies $\xi(\alpha) > 0$ for all $\alpha \in (0, 1)$, as claimed. $\square$

Let $\mu^* := (f_{21} + \lambda^* f_{22})/(f_{11} + \lambda^* f_{12})$. It is easy to see that $\mu^* > 0$. We claim

Lemma I.28. With μ^* and λ^* defined as above, $\mu^* > \lambda^*$.

Proof. $\mu^* = (f_{21} + \lambda^* f_{22})/(f_{11} + \lambda^* f_{12}) > \lambda^*$ if, and only if, $f_{21} + \lambda^* f_{22} > f_{11}\lambda^* + \lambda^{*2}f_{12}$. This holds if, and only if, $\psi(\lambda^*) > 0$, where ψ is the quadratic defined in Lemma I.26. But this is true because $\lambda^* \in (0, 1)$ and $\psi > 0$ on $(0, 1)$. $\square$

Consider, again, the equations

$$\begin{aligned}[\mathbf{PK}_1] \quad & \theta_1 x_1 + \alpha f_{11} w_{11} + \alpha f_{12} w_{12} = v_1 \\ [\mathbf{PK}_2] \quad & \theta_2 x_2 + \alpha f_{21} w_{21} + \alpha f_{22} w_{22} = v_2 \\ [\mathbf{IC}_{21}] \quad & \theta_2 x_1 + \alpha f_{21} w_{11} + \alpha f_{22} w_{12} \leq v_2\end{aligned}$$

Multiply $[\mathbf{PK}_1]$ by $t_0 = \theta_2/\theta_1 < 1$ and subtract it from $[\mathbf{IC}_{21}]$. This results in

$$\alpha w_{11}(f_{21} - t_0 f_{11}) + \alpha w_{12}(f_{22} - t_0 f_{12}) \leq v_2 - t_0 v_1$$

which we rewrite as

$$[\mathbf{IC}_{21}^*] \quad \alpha \eta_1 w_{11} + \alpha \eta_2 w_{12} \leq v_2 - t_0 v_1$$

Notice that $[\mathbf{PK}_1]$ and $[\mathbf{IC}_{21}^*]$ together imply $[\mathbf{IC}_{21}]$. To see this, multiply $[\mathbf{PK}_1]$ by $t_0 = \theta_2/\theta_1 < 1$ and add it to $[\mathbf{IC}_{21}^*]$ to obtain $[\mathbf{IC}_{21}]$.

I.3.2. Suboptimal Strategy – Part 1

Let $V^{(1)} := \{\mathbf{v} \in V_2 : \mathbf{v} = (v_1, \lambda v_1), \lambda_* < \lambda \leq \lambda^*\}$ be a cone in the domain. We shall now show that every point in $V^{(1)}$ is self-generating in the sense that there exists a feasible strategy such that given a starting point $\mathbf{v} \in V^{(1)}$, we can satisfy promise keeping and incentive compatibility while staying at $\mathbf{v}$ in all subsequent periods.

For $\mathbf{v} \in V^{(1)}$ and $s \in S$, define the menu $\zeta(\mathbf{v}, s) = (\zeta^f(\mathbf{v}, s), \zeta^c(\mathbf{v}, s))$ by $\zeta^f(\mathbf{v}, s, i) := x_i(\mathbf{v})$, where $x_i(\mathbf{v})$ is independent of s and to be solved for later, and $\zeta^c(\mathbf{v}, s, i) := \mathbf{v}$, which is independent of both s and i . When $\mathbf{v} \in V^{(1)}$ is given, we will often simply denote the flow policy by x_i .

It suffices to show the following:

Lemma I.29. Let $\mathbf{v} \in V^{(1)}$. Then, there exist $x_1, x_2 < 0$ such that $\zeta(\mathbf{v}, s)$, defined as above with these x_i , solves [\[PK₂\]](#), [\[PK₁\]](#), and [\[IC₂₁\]](#).

Proof. There exists $x_2 < 0$ such that $(x_2, \mathbf{v})$ satisfies [\[PK₂\]](#) if, and only if,

$$\alpha(f_{21} + \lambda f_{22})v_1 > \lambda v_1$$

which, in turn, holds if, and only if, $\lambda > \alpha f_{21}/(1 - \alpha f_{22}) = \lambda_*$, which is true by assumption.

Similarly, there exists $x_2 < 0$ such that $(x_2, \mathbf{v})$ satisfies [\[PK₁\]](#) if, and only if,

$$\alpha(f_{11} + \lambda f_{12})v_1 > v_1$$

which, in turn, holds if, and only if, $\lambda < (1 - \alpha f_{11})/\alpha f_{22}$, which is always true because $\lambda < 1 < (1 - \alpha f_{11})/\alpha f_{22}$.

Finally, note that [\[IC₂₁[★]\]](#) does not depend on the choice of (x_1, x_2) . Thus, it suffices to show that [\[IC₂₁[★]\]](#) is satisfied with $\mathbf{w}_1 = \mathbf{v}$. This is true if and only if

$$\alpha(\eta_1 + \lambda \eta_2)v_1 \leq (\lambda - t_0)v_1$$

Because $v_1 < 0$, this holds if, and only if, $\alpha(\eta_1 + \lambda \eta_2) \geq (\lambda - t_0)$ which is equivalent to $\lambda \leq (\alpha \eta_1 + t_0)/(1 - \alpha \eta_2) = \lambda^*$. This holds because $\mathbf{v} = (v_1, \lambda v_1) \in V^{(1)}$. This completes the proof. $\square$

I.3.3. Suboptimal Strategy — Part 2

Recall μ^* defined above in [Section I.3.1](#), and define

$$V^{(2)} := \{\mathbf{v} \in V_2 : \mathbf{v} = (v_1, \mu v_1), 0 < \mu < \mu^*\}$$

We will now show that it is possible to implement $\mathbf{v} \in V^{(2)}$ with a menu such that $\mathbf{w}_i \in V^{(1)}$ for $i = 1, 2$. More precisely:

Lemma I.30. Let $\mathbf{v} \in V^{(2)}$. Then, there exists $(x_i, \mathbf{w}_i)_{i=1,2} \ll \mathbf{0}$ that implements $\mathbf{v}$ such that $\mathbf{w}_i \in V^{(1)}$ for $i = 1, 2$.

Proof. From [Lemma I.29](#), we can implement all $\mathbf{v} \in V^{(1)}$ in this manner. Therefore, it suffices to consider $\mathbf{v} \in V^{(2)} \setminus V^{(1)}$. (Recall that $V_1 \subset V_2$ because $\lambda^* < \mu^*$, as proved in

Lemma I.28.) We can readily choose $\mathbf{w}_2 \in V^{(1)}$ and $x_2 < 0$ such that $\theta_2 x_2 + \mathbf{E}^{f_2}[\mathbf{w}_2] = v_2$, i.e., such that $[\mathbf{PK}_2]$ holds. Notice that this can be done independently of our choice of $(x_1, \mathbf{w}_1)$, as $(x_2, \mathbf{w}_2)$ does not enter in to either $[\mathbf{PK}_1]$ or $[\mathbf{IC}_{21}]$.

Let $\mathbf{v} = (v_1, v_2 = \mu v_1)$. In order to prove the lemma, we need to find $w_{11} < 0$ and $\lambda \in (\lambda_*, \lambda^*]$ such that

$$\begin{aligned}\alpha(f_{11} + \lambda f_{12})w_{11} &> v_1 \\ \alpha(\eta_1 + \lambda\eta_2)w_{11} &\leq v_2 - t_0 v_1 = (\mu - t_0)v_1\end{aligned}$$

where the first inequality is merely $[\mathbf{PK}_1]$ (with $x_1 < 0$ treated as a slack variable) while the second reflects $[\mathbf{IC}_{21}^*]$. This system has a solution if, and only if,

$$\frac{1}{\alpha(f_{11} + \lambda f_{12})}v_1 < \frac{\mu - t_0}{\alpha(\eta_1 + \lambda\eta_2)}v_1$$

which has a solution if, and only if,

$$\frac{1}{\alpha(f_{11} + \lambda f_{12})} > \frac{\mu - t_0}{\alpha(\eta_1 + \lambda\eta_2)}$$

This has a solution if, and only if,

$$\begin{aligned}\mu &< \frac{\eta_1 + \lambda\eta_2 + t_0(f_{11} + \lambda f_{12})}{(f_{11} + \lambda f_{12})} \\ &= \frac{f_{21} - t_0 f_{11} + \lambda(f_{22} - t_0 f_{12}) + t_0 f_0 - t_0 \lambda f_{12}}{(f_{11} + \lambda f_{12})} \\ &= \frac{f_{21} + \lambda f_{22}}{(f_{11} + \lambda f_{12})} \\ &\leq \mu^* \quad \text{whenever } \lambda \leq \lambda^*\end{aligned}$$

which is true by the assumption that $\mathbf{v} \in V^{(2)}$. Therefore, we can choose any $\lambda \in (\lambda_*, \lambda^*]$ for which there exists $w_{11} < 0$ so that $[\mathbf{PK}_1]$ and $[\mathbf{IC}_{21}^*]$ (or equivalently $[\mathbf{IC}_{21}]$) hold.

Notice that given $\mathbf{w}_1 \in V^{(1)}$ with the above properties, we can always find $x_1 < 0$ such that $[\mathbf{PK}_1]$ holds with equality. This completes the proof. $\square$

Let $(x_i(\mathbf{v}), \mathbf{w}_i(\mathbf{v}))_{i=1,2}$ be the solution described in Lemma I.30, for some given $\mathbf{v} \in V^{(2)}$. In this manner, for each $\mathbf{v} \in V^{(2)}$ and $s \in S$ we may define the menu $\zeta(\mathbf{v}, s)$ by $\zeta^f(\mathbf{v}, s, i) := x_i(\mathbf{v})$ and $\zeta^c(\mathbf{v}, s, i) := \mathbf{w}_i(\mathbf{v})$, both of which are independent of s .

I.3.4. Suboptimal Strategy — Part 3

Let $V^{(3)} := \{\mathbf{v} \in V_2 : \mathbf{v} = (v_1, \mu v_1), \mu^* \leq \mu < 1\}$, so that $V_2 = V^{(2)} \cup V^{(3)}$. We shall now show that it is possible to construct a contract (really, a sequence of menus) so that for

any $\mathbf{v} \in V^{(3)}$, the continuation utilities reach $V^{(1)}$ in finitely many steps, regardless of the sequence of shocks. In particular, this time can be taken to be independent of the sequence of shocks under consideration.

Let $W \subset V_2$ be a fixed subset and fix an initial $\mathbf{v} \in V_2$ and $s^{(0)} \in S$. A contract is *W-amenable* at $(\mathbf{v}, s^{(0)})$ if there is a finite natural number N (that depends on W and $(\mathbf{v}, s^{(0)})$) such that for any sequence of shocks $s^{(1)}, \dots, s^{(N)} \in S^N$, the continuation utility reaches W in at most N steps.

Lemma I.31. Let $\mathbf{v} \in V^{(3)}$ and fix $s^{(0)}$. Then, there exists a contract that is $V^{(1)}$ -amenable at $(\mathbf{v}, s^{(0)})$.

Proof. Let $\mathbf{v} = (v_1, v_2 = \mu_0 v_1) \in V^{(3)}$ where $\mu_0 \in [\mu^*, 1)$. We will show that linear inequalities derived from the constraints possess suitable solutions; any collection of such solutions will serve as the desired $V^{(1)}$ -amenable contract.

First, consider the case where $s^{(1)} = 2$. Let $\mathbf{w}_2 = (w_{21}, \lambda w_{21})$, so that we can re-write [\[PK₂\]](#) as

$$\alpha(f_{21} + \lambda f_{22})w_{21} > \mu v_1$$

(where x_2 is treated as a slack variable). It is immediate that we can choose $\lambda \in (\lambda_*, \lambda^*]$ and $w_{21} < 0$ that satisfies this inequality, and hence [\[PK₂\]](#) with some appropriately chosen $x_2 < 0$. Thus, if the first shock is $s^{(1)} = 2$, then the continuation utility $\mathbf{w}_2 \in V^{(1)}$.

Next, consider the case where $s^{(1)} = 1$. Let $\mathbf{w}_1 = (w_{11}, \mu_1 w_{11})$, so that [\[PK₁\]](#) and [\[IC₂₁[★]\]](#) can (respectively) be re-written as

$$\begin{aligned} \alpha(f_{11} + \mu_1 f_{12})w_{11} &> v_1 \\ \alpha(\eta_1 + \mu_1 \eta_2)w_{11} &\leq (\mu_0 - t_0)v_1 \end{aligned}$$

(where $x_1 < 0$ is treated as a slack variable). We can find $w_{11} < 0$ that solves the above inequalities if, and only if,

$$\frac{1}{f_{11} + \mu_1 f_{12}} > \frac{\mu_0 - t_0}{\eta_1 + \mu_1 \eta_2}$$

which, in turn, holds if, and only if,

$$\mu_1(\eta_2 - f_{12}(\mu_0 - t_0)) > f_{11}(\mu_0 - t_0) - \eta_1$$

It is easy to see that $\eta_2 - f_{12}(\mu_0 - t_0) = f_{22} - f_{12}\mu_0 > 0$. Moreover, $f_{11}(\mu_0 - t_0) - \eta_1 = f_{11}\mu_0 - f_{21}$.

Notice that $f_{11}\mu_0 - f_{21}$ can be either positive or negative. If it is negative (or zero), then we can always find $\mu_1 \in (\lambda_*, \lambda^*]$ such that $\mu_1(\eta_2 - f_{12}(\mu_0 - t_0)) > 0$, which implies that continuation utility $\mathbf{w}_1 \in V^{(1)}$ in the following period.

However, if $f_{11}\mu_0 - f_{21} > 0$, we can set $\mu_1 := (f_{11}\mu_0 - f_{21})/(f_{22} - f_{12}\mu_0)$. Then,

$$\begin{aligned}\mu_0 - \mu_1 &= \mu_0 - (f_{11}\mu_0 - f_{21})/(f_{22} - f_{12}\mu_0) \\ &= \frac{1}{f_{22} - f_{12}\mu_0} [-f_{12}\mu_0^2 + \mu_0(f_{22} - f_{11}) + f_{21}] \\ &= \frac{1}{f_{22} - f_{12}\mu_0} \psi(\mu_0) \\ &> 0\end{aligned}$$

because $\psi(t) > 0$ for $t \in (0, 1)$. Thus, $\mu_0 > \mu_1$. Proceeding iteratively, we can find a decreasing sequence where $\mu_n := (f_{11}\mu_{n-1} - f_{21})/(f_{22} - f_{12}\mu_{n-1})$. Because $\psi(\cdot)$ is a quadratic, the sequence $(\psi(\mu_n))$ lies in $[\psi(0), \psi(\mu_0)]$. Therefore,

$$\begin{aligned}\mu_{n-1} - \mu_n &= \frac{1}{f_{22} - f_{12}\mu_{n-1}} \psi(\mu_{n-1}) \\ &\geq \frac{1}{f_{22}} \min[\psi(0), \psi(\mu_0)] > 0\end{aligned}$$

so the decreasing sequence has a difference uniformly bounded away from zero. Therefore, there exists a finite N such that $\mu_N < \mu^*$. Then, if continuation utility is in $V^{(2)}$, we can use the menu defined in Lemma I.30 that then implements a transition to $V^{(1)}$ in one more step.

Any contract constructed from the solutions to these linear inequalities is one we desired, completing the proof. $\square$

I.3.5. Suboptimal Strategy — Summary

We shall now show that there exists a contract that has finite cost for the principal and is [TVC]-implementable starting from any $\mathbf{v} \in V_2$.

Proposition I.32. There exists a recursive contract ζ that (i) is [TVC]-implementable starting from any $\mathbf{v} \in V_2$ and (ii) induces a well-defined and finite-valued cost function $Q^\zeta : V_2 \times S \rightarrow \mathbb{R}$.

Proof. The proof relies on Lemmas I.29, I.30, and I.31. Starting at $(\mathbf{v}, s)$, the contract defined in Lemmas I.30 and I.31 guides continuation promised utility to $V^{(1)}$ in finitely many steps, where the finite bound is independent of the sequence of shocks. Then, upon reaching $V^{(1)}$, continuation utility stays at the same point regardless of the shock – this is Lemma I.29. Since there are only finitely many different utility levels for any starting point $(\mathbf{v}, s)$, [TVC] clearly holds and the cost to the principal of this contract is well-defined and finite, which proves the claim. $\square$

I.4. Final Steps: Proof of Theorem 3

We prove each part of the theorem in turn.

Proof of Part (a). Existence, non-emptiness, and convexity of D follow from part (a) of Proposition I.3. Part (c) of Proposition I.7 establishes that $D \supseteq V_d$ (and, hence, also non-emptiness of D). Openness and the independence properties follow from part (b) of Proposition I.25. □

Proof of Part (b). This is exactly part (c) of Proposition I.3. □

Proof of Part (c). This is exactly part (b) of Proposition I.3. □

Proof of Part (d). This was proved in SA-E.1 of BKL. □

Proof of Part (e). By part (b) of Proposition I.25, it suffices to establish the claim for the case of CARA utility. Thus, consider iterates of the operator $\mathcal{B}^\dagger$ under CARA utility. By Lemma I.12, the iterates of the operator $\mathcal{B}^\dagger$ starting with the set $\mathbb{R}_-^d$ are characterized by the sequence $(K_{t^{(n)}})$ and $K_*^\dagger$, the largest fixed point of $\mathcal{B}^\dagger$, is the limit of this sequence. That is, $K_*^\dagger = K_{t^*}$ where $t^* := \lim_{n \rightarrow \infty} t^{(n)}$.

But we have shown in Lemmas I.17 and I.18 that under either MLRP or PPR, it cannot be that $t^* > 1$. Thus, by part (a) of Lemma I.12, $t^* = 1$. Thus, $K_*^\dagger = \text{cl}V_d$. It follows from part (b) of Proposition I.8 and part (a) of Proposition I.9 that $K_* = \text{cl}V_d$. Thus, part (c) of Proposition I.9 and Proposition I.10 imply that $D = V_d$ (alternatively, this last step can be deduced from part (b) of Proposition I.25). □

Proof of Part (f). Given part (e) of Theorem 3, proved above, it suffices to show that $D^* = V_d$ under the additional assumption of CARA utility. This was established in Proposition I.32 of Section I.3. □

J. Proofs for Sections 3.2 and 3.3

This section collects proofs related to the regularity conditions from Section 3.2 and the Bellman equation from Section 3.3 of BKL. We first prove Lemma B.2 in Section J.1. In Section J.2, we present and discuss Theorem 6, which extends the treatment of the Bellman equation from Section 3.3 of BKL. We prove most of Theorem 6 in Section J.3 (part (c) of that theorem is proved separately in SA-C.3.3 of BKL).

J.1. Proof of Lemma B.2

We begin with the proof of part (a) of Lemma B.2. (Note that, if $\mathbf{v} \in V_d$, then part (b) of Lemma I.5 and Lemma I.6 together imply that $\Gamma_\circ(\mathbf{v}) \neq \emptyset$. The following proof extends the same reasoning to the potential case in which $D \not\supseteq V_d$.)

Proof of Part (a), Case (i). Let the type process satisfy FOSD. Let $\mathbf{v} \in D$ and consider any menu $(u_i, \mathbf{w}_i)_{i \in S} \in \Gamma(\mathbf{v})$. We will construct a perturbation of this menu that lies in $\Gamma_\circ(\mathbf{v})$, thus establishing nonemptiness.

Let $\ell \in S$ be given. By hypothesis, $(u_\ell, \mathbf{w}_\ell) \in \mathcal{U} \times D$ satisfies **[PK_i]** ($i = \ell$) and each of the **[IC_{ij}]** ($i > j = \ell$). Let $\varepsilon > 0$, and let $w'_{\ell k} := w_{\ell k} - \varepsilon/2^{d+1-k}$ for $k = 1, \dots, d$. Now, let $u'_\ell := u_\ell + \delta$ such that $\delta := \alpha\varepsilon \mathbf{E}^{\mathbf{f}_\ell}[(\frac{1}{2^d}, \dots, \frac{1}{2})]$. By construction, $(u'_\ell, \mathbf{w}'_\ell)$ satisfies **[PK_i]** ($i = \ell$). Note that we may take $\varepsilon > 0$ sufficiently small so that $\mathbf{w}'_\ell \in D$ and $u_\ell \in \mathcal{U}$, as both D and $\mathcal{U}$ are open sets (by part (a) of Theorem 3).

Recall from Lemma I.1 that for all $i > \ell$ we have $\psi'(\cdot, i, \ell) < 1$, which implies that $\psi(u_\ell + \delta, i, \ell) - \psi(u_\ell, i, \ell) = \int_0^\delta \psi'(u_\ell + y, i, \ell) dy < \delta$. Therefore, for all $i > \ell$, we have

$$\text{[J.1]} \quad [\psi(u_\ell + \delta, i, \ell) - \psi(u_\ell, i, \ell)] - \alpha\varepsilon \mathbf{E}^{\mathbf{f}_i}[(\frac{1}{2^d}, \dots, \frac{1}{2})] < 0$$

by definition of $\delta > 0$ and because $\mathbf{E}^{\mathbf{f}_i}[(\frac{1}{2^d}, \dots, \frac{1}{2})] > \mathbf{E}^{\mathbf{f}_\ell}[(\frac{1}{2^d}, \dots, \frac{1}{2})]$ due to the assumption that $\mathbf{f}_i$ first order stochastically dominates $\mathbf{f}_\ell$ for all $i > \ell$. Then, we have

$$\begin{aligned} \psi(u'_\ell, i, \ell) + \alpha \mathbf{E}^{\mathbf{f}_i}[\mathbf{w}'_\ell] &= \psi(u_\ell, i, \ell) + [\psi(u'_\ell, i, \ell) - \psi(u_\ell, i, \ell)] + \alpha \mathbf{E}^{\mathbf{f}_i}[\mathbf{w}_\ell - \varepsilon(\frac{1}{2^d}, \dots, \frac{1}{2})] \\ &\leq v_i + [\psi(u'_\ell, i, \ell) - \psi(u_\ell, i, \ell)] - \alpha\varepsilon \mathbf{E}^{\mathbf{f}_i}[(\frac{1}{2^d}, \dots, \frac{1}{2})] \\ &< v_i \end{aligned}$$

where the weak inequality is because $(u_\ell, \mathbf{w}_\ell)$ satisfies **[IC_{ij}]** ($i > j = \ell$), and the strict inequality is because of **[J.1]**. Thus, it is clear that $(u'_\ell, \mathbf{w}'_\ell) \in \Gamma_\circ(\mathbf{v})$. $\square$

Proof of Part (a), Case (ii). Let the transition probabilities $\{\mathbf{f}_i\}_{i=1}^d$ be affinely independent. This implies that each of the $\mathbf{f}_j$ is an extreme point of the convex hull $\text{co}(\{\mathbf{f}_i\}_{i=1}^d)$. Thus, by the Separating Hyperplane Theorem, for each $\ell \in S$ there exists a vector $\mathbf{a}_\ell \in \mathbb{R}^d$ such that $\mathbf{E}^{\mathbf{f}_i}[\mathbf{a}_\ell] > \mathbf{E}^{\mathbf{f}_\ell}[\mathbf{a}_\ell] > 0$ for all $i > \ell$. We may then adapt the proof used in Case (i) above by simply replacing the vector $(\frac{1}{2^d}, \dots, \frac{1}{2})$ with the vectors $\mathbf{a}_\ell$ when perturbing $(u_\ell, \mathbf{w}_\ell)$. $\square$

Towards the proof of part (b) of Lemma B.2, we next present a series of lemmas. Lemma J.1, stated below, does not rely on the CARA or PPR/MLRP assumptions.

Lemma J.1. Under Condition R.2, there exists a real-valued function $Q^L : D \times S \rightarrow \mathbb{R}$ such that $Q^L \leq P$ on $D \times S$.

Proof. By Lemma F.1, the first-best value function Q^* is real-valued on $\mathcal{U}^d \times S$. Let Q^L denote the restriction of Q^* to $D \times S$. Then $Q^L \leq P$ on $D \times S$. $\square$

Lemma J.2. Suppose that the agent has CARA utility and that the type process satisfies either MLRP or PPR. Then there exists a real-valued function $Q^U : D \times S \rightarrow \mathbb{R}$ such that $P \leq Q^U$ on $D \times S$.

Proof. By part (f) of Theorem 3, we know that under these conditions $D^* = D = V_d$. The finite upper bound is a consequence of the (suboptimal) [TVC]-implementable contract ζ constructed in Section I.3. By Proposition I.32, the value function induced by the contract ζ is real-valued, $Q^\zeta : D \times S \rightarrow \mathbb{R}$. Moreover, by the definition of Q^ζ and P , we have that $P \leq Q^\zeta$ on $D \times S$. Thus, let $Q^U := Q^\zeta$. $\square$

Lemma J.3, stated below, does not rely on the CARA or PPR/MLRP assumptions. It will also be used later in the proof of Theorem 6. Recall the full-information constraint correspondence $\Gamma^{\text{FB}} : \mathcal{U}^d \times S \rightarrow (\mathcal{U} \times \mathcal{U}^d)^d$ defined in SA-F of BKL.

Lemma J.3. There exists an $\eta \in \mathbb{R}_{++}$ such that, for each $\mathbf{v} \in \mathcal{U}^d$ and $(u_i, \mathbf{w}_i)_{i \in S} \in \Gamma^{\text{FB}}(\mathbf{v})$, it holds that $u_i \in (v_i, 0)$ and $w_{ij} \in (\eta v_i, 0)$ for all $i, j \in S$.

Proof. The upper bounds follow directly from Assumption DARA. The lower bounds are simple consequences of the promise keeping constraints [PK_i]. First, we have $v_i < u_i$ because $\mathbf{w}_i \in \mathcal{U}$ by Assumption DARA. Second, we have $v_i < \alpha f_{ij} w_{ij}$ because $u_i \in \mathcal{U}$ by Assumption DARA. The lemma follows by defining $\eta := 1/(\alpha \min_{i,j \in S} f_{ij})$, which is well-defined and finite by Assumption Markov. $\square$

Lemma J.4. Suppose that the agent has CARA utility. Then regularity Condition R.2 holds.

Proof. We first bound the growth rate of the principal's flow cost function. Under CARA utility, $U(c) = -e^{-c}$ and $U^{-1}(u) = -\log(-u)$. (We set the risk aversion coefficient to 1 for notational simplicity.) Let $\theta_i = e^{-\omega_i}$. Then it is easy to see that $C(u, i) = U^{-1}(u/\theta_i)$. Fix an initial promise $\mathbf{v} \in \mathcal{U}^d$, a seed type $s \in S$, and a full-information contract $\zeta \in \Xi^{\text{FB}}$. Let $(\mathbf{v}^{(t)})$ and, for each $i \in S$, $(u_i^{(t)})$ and $(\mathbf{w}_i^{(t)})$ denote the stochastic processes for promises, flow utility policies, and continuation utility policies induced by ζ initialized at $\mathbf{v}$. Fix a

path $h \in \mathcal{H}$. Recall the constant $\eta > 0$ defined in Lemma J.3. Along the path h , we have

$$\begin{aligned}
C(u_i^{(t)}(h), i) &= -\log \left(\left| \frac{v_i^{(t)}(h) - \alpha \mathbb{E}_i [\mathbf{w}_i^{(t)}(h)]}{\theta_i} \right| \right) \\
\text{[J.2]} \quad &\geq -\log \left(\frac{1}{\theta_i} \cdot \left[|v_i^{(t)}(h)| + \sum_{j=1}^d |w_{ij}^{(t)}(h)| \right] \right) \\
\text{[J.3]} \quad &\geq \log(\min_{i \in S} \theta_i) - \log \left(\eta^t \cdot \max_{i \in S} |v_i^{(0)}| \cdot [1 + d \cdot \eta] \right)
\end{aligned}$$

where the first inequality follows from the triangle inequality and monotonicity of the $\log(\cdot)$ function, and the second inequality is a t -fold iteration of the bounds from Lemma J.3. This simplifies to

$$C(u_i^{(t)}(h), i) \geq \log(\min_{i \in S} \theta_i) - t \log(\eta) - \log(\max_{i \in S} |v_i|) - \log(1 + d \cdot \eta)$$

Note that the RHS is independent of the path $h \in \mathcal{H}$ and the current type $i \in S$. Thus, taking the α -discounted sum, it is easy to see that there exist $\kappa_1 \in \mathbb{R}$ and $\kappa_2 < 0$ such that

$$\text{[J.4]} \quad \inf_{h \in \mathcal{H}} \sum_{t=0}^{\infty} \alpha^t C \left(u_{s^{(t+1)}(h)}^{(t)}(h), s^{(t+1)}(h) \right) \geq \kappa_1 + \kappa_2 \log(\max_{i \in S} |v_i|)$$

The discounted sum on the LHS is the realized cost to the principal of ζ along path h . The inequality in [J.4] gives a lower bound on *pathwise* costs that is uniform across paths and feasible full-information contracts. Note also that the constants κ_1, κ_2 are independent of both the path h and the initial condition $(\mathbf{v}, s)$. An immediate consequence of [J.4] is that

$$Q^*(\mathbf{v}, s) \geq \kappa_1 + \kappa_2 \log(\max_{i \in S} |v_i|)$$

on $\mathcal{U}^d \times S$. Let a (potentially different) path $h \in \mathcal{H}$ be given. The above inequality holds for all $(\mathbf{v}^{(t)}(h), s^{(t)}(h))$, i.e.,

$$Q^*(\mathbf{v}^{(t)}(h), s^{(t)}(h)) \geq \kappa_1 + \kappa_2 \log(\max_i |v_i^{(t)}(h)|)$$

Iterating on the same growth-rate bounds from Lemma J.3 used to derive [J.3], we may bound the RHS of the above inequality from below as follows:

$$\text{[J.5]} \quad \kappa_1 + \kappa_2 \log(\max_i |v_i^{(t)}(h)|) \geq \kappa_1 + \kappa_2 \cdot t \cdot \log(\eta) + \kappa_2 \cdot \log(\max_{i \in S} |v_i^{(0)}|)$$

The RHS of [J.5] is independent of the path $h \in \mathcal{H}$. Then, using the fact that $\alpha^t \cdot t \rightarrow 0$, we

have

$$\liminf_{t \rightarrow \infty} \alpha^t \left[\inf_{h \in \mathcal{H}} Q^*(\mathbf{v}^{(t)}(h), s^{(t)}(h)) \right] \geq \lim_{t \rightarrow \infty} \alpha^t \left[\kappa_1 + \kappa_2 \cdot t \cdot \log(\eta) + \kappa_2 \cdot \log(\max_{i \in S} |v_i^{(0)}|) \right] = 0$$

which is the desired condition. $\square$

We now collect these pieces into a proof of part (b) of Lemma B.2.

Proof of Part (b). Display [J.4] in the proof of Lemma J.4 and the fact that $\Xi \subset \Xi^{\text{FB}}$ together imply that the expected discounted sum defining P is well-defined. Lemmas J.1 and J.2 then imply that P is real-valued. This establishes that Condition R.1 is satisfied. Condition R.2 is satisfied by Lemma J.4. Condition R.3 is satisfied by case (i) of part (a) of Lemma B.2, proved above. $\square$

J.2. Bellman Equation

The following result generalizes and extends Proposition 3.2 from Subsection 3.3 of the main text, which then follows as an immediate consequence. Parts of the following theorem are used repeatedly in intermediate steps of the proof of Theorem 1 contained in SA-C of BKL.

Theorem 6. *Suppose the environment is Regular. Then the principal's value function $P : D \times S \rightarrow \mathbb{R}$ satisfies the functional equation*

$$[\text{FE}] \quad P(\mathbf{v}, s) = \min_{(u_i, \mathbf{w}_i)_{i \in S} \in \Gamma(\mathbf{v})} \sum_{i \in S} f_{si} [C(u_i, i) + \alpha P(\mathbf{w}_i, i)]$$

and, for each $s \in S$, $P(\cdot, s)$ is convex. Moreover:

- (a) P is the pointwise smallest solution to [FE] that lies pointwise above Q^* , the first-best value function.
- (b) P is strictly increasing in v_1 and non-monotone in v_i for all $i > 1$. For any convergent sequence $(\mathbf{v}^n) \subset D$ such that $\mathbf{v}^n \rightarrow \text{bd } D$, the boundary of D , we have $P(\mathbf{v}^n, s) \rightarrow +\infty$.
- (c) There exists an optimal contract ξ^* such that, for each $i \in S$, the functions $\xi^{*f}(\cdot, \cdot, i)$ and $\xi^{*c}(\cdot, \cdot, i)$ depend on $(\mathbf{v}, s)$ only through $(v_i, \dots, v_d)$.⁶
- (d) The policy correspondence derived from [FE] is nonempty-, compact-, and convex-valued and is upper hemicontinuous.
- (e) If, in addition, the environment is [TV C]-Regular, then (i) each $P(\cdot, s)$ is strictly convex and continuously differentiable, (ii) there exists a unique optimal contract ξ^* , which

⁶That is $\xi^{*f}(\mathbf{v}, s, i) = \xi^{*f}(\mathbf{v}', s', i)$ for all $(\mathbf{v}, s), (\mathbf{v}', s') \in D \times S$ such that $v_j = v'_j$ for all $j \geq i$, and similarly for ξ^{*c} .

satisfies the independence properties stated in part (c), (iii) for each $s \in S$, $\xi^*(\cdot, s)$ is a continuous function, and (iv) the induced allocation $\tilde{u}_{\xi^*}$ solves the sequence problem [SP].

The proof of Theorem 6 is in Section J.3 below. The proof makes clear precisely which of the Conditions R.1–R.4 are used to establish each part of the theorem.

While many properties described in Theorem 6 are standard, two points warrant mention. First, part (b) establishes that the value function is non-monotone in each v_i with $i > 1$. To understand this, consider the case in which $d = 2$ and the type process satisfies FOSD, so the domain $D = V_2$ (see Theorem 3 and Figure 1 of BKL). An increase in v_1 tightens both the promise keeping constraint [PK_{*i*}] ($i = 1$) and the incentive constraint [IC_{*21*}^{*}], leading to unambiguously higher costs for the principal. An increase in v_2 , on the other hand, tightens [PK_{*i*}] ($i = 2$) and adds slack to [IC_{*21*}^{*}]. The first effect increases costs, while the second effect lowers costs; depending on the current state $(\mathbf{v}, s)$, either of these effects can dominate. When $v_1 \approx 0$ (i.e., $\mathbf{v}$ is close to the upper boundary of V_2), the first effect dominates and P is increasing in v_2 . When $\mathbf{v}$ is close enough to the diagonal (i.e., the lower boundary of V_2), the second effect dominates and P is decreasing in v_2 .

Second, part (c) states that the optimal contract does not depend on all components of the state variable $(\mathbf{v}, s)$. The independence of the policy functions on the previous report $s \in S$ follows directly from the interim nature of the promised utility state variable. Because the constraint set $\Gamma(\mathbf{v})$ does not depend on s , it is easy to see that the optimization problem in [FE] may be decomposed into d separate problems, each of which conditions only on the agent's *current* report. (This property is used in SA-C.3.3 of BKL as part of the proof of Theorem 1.) Thus, $\mathbf{v} \in D$ is a “sufficient statistic” for $(\mathbf{v}, s)$ for the purpose of computing optimal flow and continuation utilities. Of course, the agent's previous report $s \in S$ is essential for determining the principal's value at each history, and is thus a necessary component of the principal's state variable. The fact that the policy functions for type $i \in S$ depend on the current promised utility $\mathbf{v}$ *only* through its components $(v_i, v_{i+1}, \dots, v_d)$ is a consequence of Assumption NHB, which implies that u_i and $\mathbf{w}_i$ only appear in those incentive constraints that deter deviations by *higher* types.

J.3. Proof of Theorem 6

We will establish the relevant properties through a series of lemmas, and collect them into a proof of the theorem at the end.

Let $\overline{\mathbb{R}}$ denote the extended reals, and let $\overline{\mathbb{R}}^{D \times S}$ denote the space of functions $f :$

$D \times S \rightarrow \overline{\mathbb{R}}$. Define the Bellman operator $T : \overline{\mathbb{R}}^{D \times S} \rightarrow \overline{\mathbb{R}}^{D \times S}$ by

$$[\mathbf{T}] \quad TQ(\mathbf{v}, s) := \inf_{(u_i, \mathbf{w}_i)_{i \in S} \in \Gamma(\mathbf{v})} \sum_{i=1}^d f_{si} [C(u_i, i) + \alpha Q(\mathbf{w}_i, i)]$$

Lemma J.5. Suppose that Condition [R.1](#) holds. Then the principal's value function, P , satisfies the Bellman equation

$$[\mathbf{BE}] \quad P = TP$$

Proof. This follows from standard Principle of Optimality arguments, e.g., a straightforward adaptation of Theorem 4.2 of Stokey, Lucas, and Prescott (1989, p. 71). $\square$

Lemma J.6. Suppose that Conditions [R.1](#) and [R.2](#) hold. Then P is the smallest fixed point of the Bellman operator $[\mathbf{T}]$ in the order interval $[Q^*, \infty]$ (endowed with the pointwise order).

Proof. By Lemma [J.5](#), it suffices to show that any fixed point $Q = TQ$ with $Q \geq Q^*$ satisfies $Q \geq P$. Because $P \geq Q^*$ by definition of these functions, Condition [R.2](#) implies that the appropriate analogue of condition 2.17 in Kamihigashi (2014) holds (with his upper bound $\bar{v}$ replaced by the lower bound Q^* and the sign of the limit term reversed, as is appropriate given our focus on minimization and his focus on maximization). Then, an elementary adaptation of Proposition 2.1 in Kamihigashi (2014) establishes the claim. $\square$

Lemma J.7. Suppose that Condition [R.1](#) holds. Then $P(\cdot, s)$ is convex and continuous on D for each $s \in S$.

Proof. By Theorem [3\(b\)](#), the constraint correspondence $\Gamma(\cdot)$ has a convex graph. In addition, the return functions $C(\cdot, s)$ are convex. It then follows from the definition of [\[RP\]](#) and standard arguments that $P(\cdot, s)$ is convex. Continuity of $P(\cdot, s)$ follows because D is open by Theorem [3\(a\)](#). $\square$

We will use the following notation in the coming proofs. For each menu $(\mathbf{u}, \mathbf{w}) := (u_i, \mathbf{w}_i)_{i \in S} \in (\mathcal{U} \times D)^d$ and $s \in S$, define

$$H(P, (\mathbf{u}, \mathbf{w}), s) := \sum_{i=1}^d f_{si} [C(u_i, i) + \alpha P(\mathbf{w}_i, i)]$$

Thus, in particular, the Bellman equation [\[BE\]](#) may be written as $P(\mathbf{v}, s) = \inf_{(\mathbf{u}, \mathbf{w}) \in \Gamma(\mathbf{v})} H(P, (\mathbf{u}, \mathbf{w}), s)$. Define the *policy correspondence* by $\Gamma^* : D \times S \rightrightarrows (\mathcal{U} \times D)^d$ by

$$[\mathbf{J.6}] \quad \Gamma^*(\mathbf{v}, s) := \arg \min_{(\mathbf{u}, \mathbf{w}) \in \Gamma(\mathbf{v})} H(P, (\mathbf{u}, \mathbf{w}), s)$$

Lemma J.8. Suppose that Conditions [R.1](#) and [R.2](#) hold. For any convergent sequence $(\mathbf{v}^n)_{n=0}^\infty \subset D$ with limit point $\mathbf{v}^* \in \text{bd } D$, the boundary of D , we have $P(\mathbf{v}^n, s) \rightarrow +\infty$ for each $s \in S$.

Proof. Let $s \in S$ be fixed throughout. Let $(\mathbf{u}^n, \mathbf{w}^n) := (u_i^n, \mathbf{w}_i^n)_{i \in S}$ denote a sequence of menus with the following properties. First, $(\mathbf{u}^n, \mathbf{w}^n) \in \Gamma(\mathbf{v}^n)$ for all $n \in \mathbb{N}$. Now, because the sequence $(\mathbf{v}^n)$ converges and is thus bounded, it follows from [Lemma J.3](#) that there exists a bounded set $R \subset (\mathcal{U} \times D)^d$ such that $(\mathbf{u}^n, \mathbf{w}^n) \in R$ for all $n \in \mathbb{N}$. Second, suppose that there exists some $\varepsilon > 0$ such that each $(\mathbf{u}^n, \mathbf{w}^n)$ is ε -optimal, i.e.,

$$[\text{J.7}] \quad H^n := H(P, (\mathbf{u}^n, \mathbf{w}^n), s) \leq P(\mathbf{v}, s) + \varepsilon$$

for all $n \in \mathbb{N}$. Such an $\varepsilon > 0$ and a corresponding sequence of menus exists by virtue of [Lemma J.5](#).

We claim that $H^n \rightarrow +\infty$, which implies that $P(\mathbf{v}^n, s) \rightarrow +\infty$. This claim follows from two facts. First, we assert that there exists some $K \in \mathbb{R}$ such that

$$\inf_{i \in S, (\mathbf{u}, \mathbf{w}) \in R} P(\mathbf{w}_i, i) \geq K$$

To see this, observe that $\inf_{i \in S, (\mathbf{u}, \mathbf{w}) \in \text{cl } R} Q^*(\mathbf{w}_i, i) > -\infty$ by [Lemma F.1](#), where Q^* is the first-best value function. The first fact then follows from the fact that $P \geq Q^*$ on $D \times S$. Second, we assert that $\mathbf{u}^n \rightarrow \mathbf{0}$. This follows from the structure of the domain D — see, in particular, the statements and proofs of [Lemma I.23](#) and [Proposition I.24](#). It is then easy to see that the claim that $H^n \rightarrow +\infty$ follows, as we have

$$H^n \geq \sum_{i=1}^d f_{si} C(u_i^n, i) + \alpha K$$

by the first fact, and $C(u_i^n, i) \rightarrow +\infty$ for all $i \in S$ by the second fact. This completes the proof. $\square$

Following [Kreps \(1977\)](#), we say that a recursive contract ξ is *conserving* if it attains the infimum in [\[FE\]](#) for each $(\mathbf{v}, s) \in D \times S$. (Note that, by definition, if ξ is conserving then it must be feasible.) It is easy to see that if ξ is optimal, then it must be conserving (though, the converse requires additional argument).

A *Kakutani correspondence* is a correspondence that is nonempty-, compact-, and convex-valued and is continuous.

Lemma J.9. Suppose that Conditions [R.1](#) and [R.2](#) hold. For each $(\mathbf{v}, s) \in D \times S$, the infimum in [\[BE\]](#) is attained. That is, there exists a conserving recursive contract ξ^* . Moreover,

the policy correspondence Γ^* is nonempty-, convex-, and compact-valued and is upper hemi-continuous.

Proof. We will construct a Kakutani correspondence and show it is without loss of optimality to restrict attention to menus that lie in this correspondence. Let $(\mathbf{v}, s) \in D \times S$ be given. Define

$$L(\mathbf{v}, s) := \{(\mathbf{u}, \mathbf{w}) \in \Gamma(\mathbf{v}) : H(P, (\mathbf{u}, \mathbf{w})) \leq P(\mathbf{v}, s) + 1\}$$

Define $K(\mathbf{v}, s) := \text{cl } L(\mathbf{v}, s)$, the closure of $L(\mathbf{v}, s)$. It follows from Lemma J.3 that $\Gamma(\mathbf{v})$ is a bounded set. Thus, it follows that $L(\mathbf{v}, s)$ is bounded and thus that $K(\mathbf{v}, s)$ is compact. It is also true that $K(\mathbf{v}, s) \neq \emptyset$ because P satisfies the Bellman equation (Lemma J.5), and thus there exist ε -optimal menus for each $\varepsilon > 0$. Moreover, it is easy to see that the function $H(P, \cdot) : \Gamma(\mathbf{v}) \rightarrow \mathbb{R}$ is convex, and thus that $K(\mathbf{v}, s)$ is a convex set (since $\Gamma(\mathbf{v})$ is convex by part (b) of Theorem 3).

We assert that the implied correspondence

$$\begin{aligned} K : D \times S &\rightrightarrows (\mathcal{U} \times D)^d \\ (\mathbf{v}, s) &\mapsto K(\mathbf{v}, s) \end{aligned}$$

is a well-defined Kakutani correspondence. To show that it is a well-defined correspondence, we must show that $K(\mathbf{v}, s) \in (\mathcal{U} \times D)^d$ for all $(\mathbf{v}, s) \in D \times S$. Clearly we must have $K(\mathbf{v}, s) \subseteq \text{cl } \Gamma(\mathbf{v})$. If $K(\mathbf{v}, s) \cap (\text{cl } \Gamma(\mathbf{v}) \setminus \Gamma(\mathbf{v})) \neq \emptyset$, then there exists some sequence $((\mathbf{u}^n, \mathbf{w}^n))_{n=0}^\infty \subseteq L(\mathbf{v}, s)$ such that $(\mathbf{u}^n, \mathbf{w}^n) \rightarrow \text{bd } (\mathcal{U} \times D)^d$. (It is easy to see that $(\text{cl } \Gamma(\mathbf{v}) \setminus \Gamma(\mathbf{v})) \subseteq \text{bd } (\mathcal{U} \times D)^d$.) By the same arguments as in the proof of Lemma J.8, $H(P, (\mathbf{u}^n, \mathbf{w}^n)) \rightarrow +\infty$, which contradicts the definition of $L(\mathbf{v}, s)$ and finiteness of $P(\mathbf{v}, s)$ (Condition R.1). Thus, K is a well-defined correspondence. To show that it is Kakutani, it suffices to show that it is continuous. But this is easy to see, as $\Gamma(\cdot)$ is a continuous correspondence, P is a continuous function (Lemma J.7), and each $H(P, \cdot, s)$ is a continuous function on $(\mathcal{U} \times D)^d$ by virtue of the continuity of the $C(\cdot, i)$ and P .

It follows that the argmin correspondence M^* defined by

$$M^*(\mathbf{v}, s) := \arg \min_{(\mathbf{u}, \mathbf{w}) \in K(\mathbf{v}, s)} H(P, (\mathbf{u}, \mathbf{w}), s)$$

is continuous and nonempty- and compact-valued by Berge's Theorem, and convex-valued because K is Kakutani and $H(P, \cdot, s)$ is convex. We showed above that $K(\mathbf{v}, s) \subseteq \Gamma(\mathbf{v})$. It is also easy to see that the policy correspondence Γ^* (defined in [J.6]) satisfies $\Gamma^*(\mathbf{v}, s) \subseteq K(\mathbf{v}, s)$. Thus, it follows that $M^*(\mathbf{v}, s) = \Gamma^*(\mathbf{v}, s)$ for all $(\mathbf{v}, s) \in D \times S$. This completes the proof. $\square$

Lemma J.10. Suppose that Conditions [R.1](#) and [R.2](#) hold. Any conserving contract ξ^* is optimal, i.e., is optimal in the principal's recursive problem [\[RP\]](#). Hence, there exists a optimal contract.

Proof. Lemma [J.9](#) shows that a conserving contract exists, so it suffices to show that any such contract is optimal for the principal. But optimality is a direct consequence of Condition [R.2](#) and an adaptation of Theorem 4.5 in Stokey, Lucas, and Prescott (1989). In particular, because P satisfies the Bellman equation [\[BE\]](#) (Lemma [J.5](#)) and the infimum is attained at each step (Lemma [J.9](#)), induction gives

$$P(\mathbf{v}, s) = \mathbf{E} \left[\sum_{t=0}^T \alpha^t C(u^{(t)}, s^{(t+1)}) \middle| s^{(0)} = s \right] + \mathbf{E} \left[\alpha^{T+1} P(\mathbf{v}^{(T+1)}, s^{(T+1)}) \middle| s^{(0)} = s \right]$$

where $(u^{(t)})_{t=0}^\infty$ is the flow utility process induced by conserving recursive contract (i.e., some measurable selection from the policy correspondence). Now, Condition [R.2](#) implies that

$$\liminf_{t \rightarrow \infty} \alpha^t \left[\inf_{h \in \mathcal{H}} P(\mathbf{v}^{(t)}(h), s^{(t)}(h)) \right] \geq 0$$

because $P \geq Q^*$ on $D \times S$. Thus, for some sufficiently large T , there exists a constant $K \in \mathbb{R}$ such that for all $t \geq T$,

$$\alpha^t \left[\inf_{h \in \mathcal{H}} P(\mathbf{v}^{(t)}(h), s^{(t)}(h)) \right] + K \geq 0$$

Thus, we may apply Fatou's Lemma to get

$$\begin{aligned} 0 &\leq \mathbf{E} \left[\liminf_{T \rightarrow \infty} \alpha^{T+1} P(\mathbf{v}^{(T+1)}, s^{(T+1)}) \middle| s^{(0)} = s \right] \\ &\leq \liminf_{T \rightarrow \infty} \mathbf{E} \left[\alpha^{T+1} P(\mathbf{v}^{(T+1)}, s^{(T+1)}) \middle| s^{(0)} = s \right] \end{aligned}$$

Combining the above displays, we get

$$\text{[J.8]} \quad P(\mathbf{v}, s) \geq \liminf_{T \rightarrow \infty} \mathbf{E} \left[\sum_{t=0}^T \alpha^t C(u^{(t)}, s^{(t+1)}) \middle| s^{(0)} = s \right]$$

By Condition [R.1](#) (which requires that the double integral be well-defined and finite), we

have

$$\begin{aligned} \mathbf{E} \left[\sum_{t=0}^{\infty} \alpha^t \max \{0, C(u^{(t)}, s^{(t+1)})\} \middle| s^{(0)} = s \right] &< +\infty \\ \mathbf{E} \left[\sum_{t=0}^{\infty} \alpha^t \max \{0, -C(u^{(t)}, s^{(t+1)})\} \middle| s^{(0)} = s \right] &< +\infty \end{aligned}$$

and thus,

$$\mathbf{E} \left[\sum_{t=0}^{\infty} \alpha^t |C(u^{(t)}, s^{(t+1)})| \middle| s^{(0)} = s \right] < +\infty$$

Then, by Fubini's Theorem,

$$\lim_{T \rightarrow \infty} \mathbf{E} \left[\sum_{t=0}^T \alpha^t C(u^{(t)}, s^{(t+1)}) \middle| s^{(0)} = s \right] = \mathbf{E} \left[\sum_{t=0}^{\infty} \alpha^t C(u^{(t)}, s^{(t+1)}) \middle| s^{(0)} = s \right]$$

Then it follows from [J.8] that $P(\mathbf{v}, s) \geq \mathbf{E} \left[\sum_{t=0}^{\infty} \alpha^t C(u^{(t)}, s^{(t+1)}) \middle| s^{(0)} = s \right]$. The opposite inequality holds by the definition of P in [RP]. Thus, $P(\mathbf{v}, s) = \mathbf{E} \left[\sum_{t=0}^{\infty} \alpha^t C(u^{(t)}, s^{(t+1)}) \middle| s^{(0)} = s \right]$ and hence the conserving contract is optimal. $\square$

Lemma J.11. Suppose that Conditions R.1 and R.5 hold. Then $P(\cdot, s)$ is continuously differentiable on D for each $s \in S$.

Proof. Under Condition R.1, Lemma J.7 states that $P(\cdot, s)$ is convex. Under Condition R.5, $P(\cdot, s)$ is differentiable (on the open set D , by Theorem 3(a)). As differentiable convex functions are continuously differentiable, the lemma follows. $\square$

Lemma J.12. Suppose that Conditions R.1 and R.4 hold. Then $P(\cdot, s)$ is *strictly* convex for each $s \in S$.

Proof. Let ξ^* be a [TVC]-implementable optimal contract, the existence of which is guaranteed by Condition R.4. Let $\mathbf{v}, \mathbf{v}' \in D$ be given such that $\mathbf{v} \neq \mathbf{v}'$. For fixed $\eta \in (0, 1)$, define $\mathbf{v}'' := \eta \mathbf{v} + (1 - \eta) \mathbf{v}'$. Let $(\mathbf{u}^{(t)}, \mathbf{w}^{(t)})_{t=0}^{\infty}$ and $(\mathbf{u}'^{(t)}, \mathbf{w}'^{(t)})_{t=0}^{\infty}$ denote the stochastic processes of menus generated by ξ^* starting from $\mathbf{v}$ and $\mathbf{v}'$, respectively. For each $t \in \mathbb{N}$, let

$$(\mathbf{u}''^{(t)}, \mathbf{w}''^{(t)}) := \eta(\mathbf{u}^{(t)}, \mathbf{w}^{(t)}) + (1 - \eta)(\mathbf{u}'^{(t)}, \mathbf{w}'^{(t)})$$

Thus, the process $(\mathbf{u}''^{(t)}, \mathbf{w}''^{(t)})_{t=0}^{\infty}$ denotes the *history-wise* convex combination of the processes $(\mathbf{u}^{(t)}, \mathbf{w}^{(t)})_{t=0}^{\infty}$ and $(\mathbf{u}'^{(t)}, \mathbf{w}'^{(t)})_{t=0}^{\infty}$. Because ξ^* is feasible and delivers promises (i.e., satisfies [DP], by virtue of Lemma A.1) and because the constraint correspondence Γ

has convex graph (by part (b) of Theorem 3), it follows that the process $(\mathbf{u}''^{(t)}, \mathbf{w}''^{(t)})_{t=0}^\infty$ is feasible and delivers promises starting from $\mathbf{v}'' \in D$.

We will establish strict convexity by expanding the right-hand side of [FE]. It follows from the facts (i) that $\mathbf{v} \neq \mathbf{v}'$, (ii) that ξ^* delivers promises, and (iii) that the type process is fully connected (by Assumption Markov) that there exists some $T \in \mathbb{N}$ such that $\mathbf{P}(\mathbf{u}^{(T)} \neq \mathbf{u}'^{(T)}) > 0$. Now, expanding the right-hand side of [FE] for T periods, we have

$$\begin{aligned}
& \eta P(\mathbf{v}, s) + (1 - \eta) P(\mathbf{v}', s) \\
&= \sum_{t=0}^T \alpha^t \mathbf{E} \left[\eta C(u_{s^{(t+1)}}^{(t)}, s^{(t+1)}) + (1 - \eta) C(u_{s^{(t+1)}}'^{(t)}, s^{(t+1)}) \middle| s^{(0)} = s \right] \\
&+ \alpha^T \mathbf{E} \left[\eta P(\mathbf{w}_{s^{(T)}}'^{(T+1)}, s^{(T+1)}) + (1 - \eta) P(\mathbf{w}_{s^{(T+1)}}'^{(T)}, s^{(T+1)}) \middle| s^{(0)} = s \right] \\
&> \sum_{t=0}^T \alpha^t \mathbf{E} \left[C(u_{s^{(t+1)}}''^{(t)}, s^{(t+1)}) \middle| s^{(0)} = s \right] + \alpha^T \mathbf{E} \left[P(\mathbf{w}_{s^{(T+1)}}''^{(T)}, s^{(T+1)}) \middle| s^{(0)} = s \right] \\
&\geq P(\mathbf{v}'', s)
\end{aligned}$$

The strict inequality follows from convexity of $P(\cdot, s)$ (established in Lemma J.7), strict convexity of the $C(\cdot, i)$, and the fact that $\mathbf{P}(\mathbf{u}^{(T)} \neq \mathbf{u}'^{(T)}) > 0$. The weak inequality follows from the definition of $P(\mathbf{v}'', s)$. This establishes strict convexity. $\square$

Lemma J.13. Suppose that Conditions R.1 and R.2 hold. For each $s \in S$, the value function $P(\cdot, s)$ is (a) strictly increasing in the component v_1 and (b) non-monotone in the component v_j for all $j = 2, \dots, d$.

Proof. Consider part (a). Let $\mathbf{v} \in D$ and $s \in S$ be given, and define $\mathbf{v}' := (v_1 - \varepsilon, v_2, \dots, v_d)$ for some sufficiently small $\varepsilon > 0$ so that $\mathbf{v}' \in D$. Let $(u_i^*, \mathbf{w}_i^*)_{i \in S}$ be an optimal menu at $(\mathbf{v}, s)$ (such a menu exists by Lemma J.9). It is easy to see that the menu $(u_i', \mathbf{w}_i')_{i \in S}$ defined by $u_1' := u_1^* - \varepsilon$, $u_j' := u_j^*$ for all $j \geq 2$, and $\mathbf{w}_i' := \mathbf{w}_i^*$ for all $i \in S$ satisfies $(u_i', \mathbf{w}_i')_{i \in S} \in \Gamma(\mathbf{v}')$. It satisfies the promise keeping constraints $[\mathbf{PK}_i]$ by construction, and adds (potentially zero) slack to all of the incentive constraints $[\mathbf{IC}_{ij}]$ ($i > j$). (Note that v_1 does not appear in any of the incentive constraints.) This menu is strictly less costly than the optimal menu at $\mathbf{v}$, so by revealed preference we have $P(\mathbf{v}', s) < P(\mathbf{v}, s)$. This proves part (a).

Part (b) follows immediately from part (a) of Theorem 3, Proposition I.11, and Lemma J.8 above. In particular, those first two results imply that $D \subseteq K_{\mathbf{t}(1)}$ (the cone $K_{\mathbf{t}(1)}$ is constructed in Proposition I.11). It is easy to see from the structure of that set that, for any $j \geq 2$ and fixed $\mathbf{v}_{-j} := (v_1, \dots, v_{j-1}, v_{j+1}, \dots, v_d)$, the set $\{y \in \mathcal{U} : (y, \mathbf{v}_{-j}) \in D\}$ is bounded. Non-monotonicity in the j th component then follows from Lemma J.8. $\square$

We may now collect the pieces above to prove the theorem.

Proof of Theorem 6. Lemmas J.5 and J.9 together imply that P satisfies [FE]. Convexity follows from Lemma J.7. Part (a) of the theorem follows from Lemma J.6. Part (b) of the theorem follows from Lemmas J.8 and J.13. For part (c), existence of an optimal contract follows from Lemma J.10. The independence properties follow from part (a) of Lemma C.12 in SA-C.3.3 of BKL. Part (d) follows from Lemma J.9.

Finally, consider part (e). Point (i), strict convexity and continuous differentiability, follows from Lemmas J.12 and J.11. Given strict convexity and Lemma J.9, the policy correspondence Γ^* is singleton-valued, so there exists a unique conserving contract and hence at most one optimal contract. By Lemma J.10, it is indeed the unique optimal contract, establishing point (ii). Point (iii), continuity, follows from uniqueness and the upper hemi-continuity in Lemma J.9. Point (iv) follows from Lemma A.1. $\square$

K. Details for Section 5.2

This section presents facts that support Properties (a)–(e) stated in Section 5.2 of BKL. We consider a setting with binary types ($d = 2$), FOSD transitions ($f_{11} \geq f_{21}$, $f_{22} \geq f_{12}$), and CARA utility: $U(c) = -e^{-\rho c}$ for some $\rho > 0$. It is convenient to define the marginal utility parameter $\theta_i := e^{-\rho \omega_i}$ for each $i \in S$, so that $U(\omega_i + c_i) = \theta_i U(c_i)$ and $\psi(u, i, j) = (\theta_i/\theta_j)u$. Under these assumptions, the environment is Regular and, by Theorem 3(f), the domain satisfies $D = D^* = V_2$, where D^* is defined in [B.1] of Appendix B.1. We further assume that Conditions R.4 and R.5 hold, so the environment is [TVC]-Regular. We let ξ^* denote the optimal contract (which exists and is unique by Proposition 3.2), and let $u_i(\mathbf{v}, s) := \xi^{*f}(\mathbf{v}, s, i)$ and $w_i(\mathbf{v}, s) := \xi^{*c}(\mathbf{v}, s, i)$ denote the policy functions.

The first fact formalizes the claim in Section 5.2 that optimal contracts are HD1, from which Property (a) follows immediately.

Fact 1. *For any persistence level, the optimal contract is homogeneous of degree one (HD1) in $\mathbf{v}$, and the solutions to the efficiency problems [Eff_i] are HD1 in v .*

The next two facts establish basic facts about the dynamics of promised utility under i.i.d. and persistent shocks, respectively.

Fact 2. *When types are iid:*

- (a) *The rays E_1 and E_2 coincide. (Call the single ray E .)*
- (b) *The policy functions satisfy $w_1(\mathbf{v}, s) \in E$ and $w_2(\mathbf{v}, s) \in E$ for all $(\mathbf{v}, s) \in V_2 \times S$.*
- (c) *Given $\mathbf{v} \in E$, $w_1(\mathbf{v}, s) \ll \mathbf{v} \ll w_2(\mathbf{v}, s)$.*

Fact 3(b) yields Property (b), from which Properties (c) and (e) immediately follows. Fact 2 yields the i.i.d. portion of Property (d). Parts (a) and (d) of Fact 3 yield the persistent portion of Property (d).

Fact 3. *With positive persistence (i.e., when $f_{11} - f_{21} \equiv f_{22} - f_{12} > 0$):*

- (a) *The ray E_1 lies strictly above the ray E_2 .*
- (b) *The policy function $\mathbf{w}_2$ satisfies $\mathbf{w}_2(\mathbf{v}, s) \in E_2$ for all $(\mathbf{v}, s) \in V_2 \times S$.*
- (c) *Given $\mathbf{v} \in E_2$, $\mathbf{w}_2(\mathbf{v}, s) \gg \mathbf{v}$.*
- (d) *On the optimal path, the policy function $\mathbf{w}_1(\mathbf{v}^{(t)}, s^{(t)})$ lies strictly below the ray E_1 .*

The final fact restates the result from Thomas and Worrall (1990) that, with i.i.d. types and CARA utility, the optimal contract is *order-independent*. This fact is referenced at the end of [Section 5.2](#).

Fact 4. *Suppose that types are iid. Fix $\mathbf{v}^{(0)} \in E$, and let h be a history of realized types and $h^t = (s^1, \dots, s^t)$ be its length- t restriction. Then $\mathbf{v}^{(t)}(h)$ depends on h^t only through the **number** of high and low reports.*

In [Sections K.1](#) and [K.2](#) below, we present and prove a series of intermediate lemmas. In [Section K.3](#), we consolidate these lemmas into proofs of the above Facts 1–4.

K.1. General Homogeneity Properties

We begin with homogeneity properties of the value function and optimal contract. These not require any assumptions on the number of types $d \in \mathbb{N}$ or the type process beyond Assumption [Markov](#).

Lemma K.1. Suppose the environment is [\[TVC\]](#)-regular. For all $a \in \mathbb{R}$ and $s \in S$:

- (a) The principal's value function satisfies the homogeneity property

$$[\text{K.1}] \quad P(e^{-a}\mathbf{v}, s) = P(\mathbf{v}, s) + \frac{a}{1 - \alpha}$$

- (b) For fixed $s \in S$, the optimal contract is homogeneous of degree one in $\mathbf{v}$

$$[\text{K.2}] \quad \xi(e^{-a}\mathbf{v}, s) = e^{-a}\xi(\mathbf{v}, s)$$

Proof. Part (a) is immediate from the facts that (i) the return function $C(x) := U^{-1}(x)$ satisfies

$$[\text{K.3}] \quad C(e^{-a}x) = -\log(-e^{-a}x) = -\log(-x) - \log(e^{-a}) = C(x) + a$$

and (ii) that the graph of the constraint correspondence $\Gamma(\cdot)$ is a convex cone by part (c) of Lemma [I.2](#). Part (b) follows from part (a) and point (ii) in the previous sentence (recall that the optimal contract is unique under [\[TVC\]](#)-regularity). $\square$

We also have the following useful homogeneity property for the derivatives of the value function:

Lemma K.2. Suppose the environment is [TVC]-regular. For all $a \in \mathbb{R}$ and $s \in S$:

(a) The partial derivatives $P_i(\cdot, i)$ are homogeneous of degree -1 , i.e., for each $i \in S$,

$$[\text{K.4}] \quad P_i(e^{-a}\mathbf{v}, s) = e^a P_i(\mathbf{v}, s)$$

(b) The directional derivative $D_1 P(\cdot, s)$ is homogeneous of degree -1 , i.e.,

$$[\text{K.5}] \quad D_1 P(e^{-a}\mathbf{v}, s) = e^a D_1 P(\mathbf{v}, s)$$

(c) For each $i \in S$, the efficiency set E_i (defined in SA-C.3.1) are rays in D .

(d) If types are iid, there exists a ray E in D such that $E = E_1 = \dots = E_d$.

Proof. For part (a), let $\varepsilon > 0$ and $i \in S$. Let $\hat{\mathbf{e}}_i \in \mathbb{R}^d$ denote the i th standard basis vector. We have:

$$\begin{aligned} P_i(e^{-a}\mathbf{v}, s) &= \lim_{\varepsilon \rightarrow 0} \frac{P(e^{-a}\mathbf{v} + \varepsilon \hat{\mathbf{e}}_i, s) - P(e^{-a}\mathbf{v}, s)}{\varepsilon} \\ &= \lim_{\varepsilon \rightarrow 0} \frac{P(e^{-a}[\mathbf{v} + e^a \varepsilon \hat{\mathbf{e}}_i], s) - P(e^{-a}\mathbf{v}, s)}{\varepsilon} \\ &= \lim_{\varepsilon \rightarrow 0} \frac{P(\mathbf{v} + e^a \varepsilon \hat{\mathbf{e}}_i, s) - P(\mathbf{v}, s)}{\varepsilon} && \text{by part (a) of Lemma K.1} \\ &= e^a \cdot \lim_{\varepsilon \rightarrow 0} \left[\frac{P(\mathbf{v} + e^a \varepsilon \hat{\mathbf{e}}_i, s) - P(\mathbf{v}, s)}{e^a \varepsilon} \right] = e^a P_1(\mathbf{v}, s) \end{aligned}$$

which proves part (a). Part (b) is an immediate corollary of part (a). For part (c), note that $[\tilde{\mathbf{E}}_i]$ together with part (a) imply that E_i must be a cone. That it must be a ray then follows from, e.g., Lemma C.5. Part (d) follows immediately from part (c) and examination of the optimality conditions in SA-C.1. $\square$

Lemma K.3. Suppose the environment is [TVC]-regular, and that the type process is iid. Let $(\mathbf{v}^{(t)})$ denote the induced promise process induced by the optimal contract. Suppose that the initial $\mathbf{v}^{(0)} \in E$, where E is defined in part (d) of Lemma K.2. Then there exist numbers $b_i \in \mathbb{R}$ for each $i \in S$ such that

$$\mathbf{v}^{(t)} = \left(\prod_{i=1}^d b_i^{N_i^{(t)}} \right) \cdot \mathbf{v}^{(0)}$$

where $(N_i^{(t)})_{t=0}^\infty$ is the counting process defined by

$$N_i^{(t)} := \left| \{T \in \{1, \dots, t\} : s^{(T)} = i\} \right|$$

Proof. By Remark C.7, the optimal contract maps E to E , i.e., $\mathbf{w}_i(\mathbf{v}, s) \in E$ for all $i \in S$ and $(\mathbf{v}, s) \in E \times S$. Given this, the lemma is an immediate consequence of part (b) of

Lemma K.1 and part (d) of Lemma K.2. □

K.2. Facts about Multipliers in the $d = 2$ Case

Lemma K.4. Suppose the environment is [TVC]-regular, and that $d = 2$. If $\mathbf{v} \in E_s$, then $\mu_{21}(\mathbf{v}, s) > 0$ for each $s \in S$. Thus, if the type process is iid and $\mathbf{v} \in E$, then $\mu_{21}(\mathbf{v}, s) > 0$ for each $s \in S$.

Proof. The first statement is a corollary of Lemmas C.16 and C.17. The second statement follows from the first and part (d) of Lemma K.2. □

Lemma K.5. Suppose the environment is [TVC]-regular, and that $d = 2$. If $\mathbf{v} \in E_s$, then $u_2(\mathbf{v}, s) > u_1(\mathbf{v}, s)$ for each $s \in S$. Thus, if the type process is iid and $\mathbf{v} \in E$, then $u_2(\mathbf{v}, s) > u_1(\mathbf{v}, s)$ for each $s \in S$.

Proof. In the $d = 2$ case, the optimality condition [FOC u_i] for $i = 2$ reduces to

$$C'(u_2(\mathbf{v}, s), 2) = \frac{\lambda_2(\mathbf{v}, s)}{f_{s2}} + \frac{\mu_{21}(\mathbf{v}, s)}{f_{s2}}$$

and [FOC u_i] for $i = 1$ reduces to

$$C'(u_1(\mathbf{v}, s), 1) = \frac{\lambda_1(\mathbf{v}, s)}{f_{s1}} - \frac{\mu_{21}(\mathbf{v}, s)}{f_{s1}} \cdot \frac{\theta_2}{\theta_1}$$

Since $\lambda_i(\mathbf{v}, s) = P_i(\mathbf{v}, s)$ by the envelope condition [Env $_i$], and since $\mathbf{v} \in E_s$, it follows from [E $_i$] that $\lambda_2(\mathbf{v}, s)/f_{s2} = \lambda_1(\mathbf{v}, s)/f_{s1}$. Thus, combining the above displays gives

$$C'(u_2(\mathbf{v}, s), 2) - C'(u_1(\mathbf{v}, s), 1) = \left(\frac{f_{s1}}{f_{s2}} + \frac{\theta_2}{\theta_1} \right) \cdot \frac{\mu_{21}(\mathbf{v}, s)}{f_{s1}}$$

Thus $C'(u_2(\mathbf{v}, s), 2) - C'(u_1(\mathbf{v}, s), 1) > 0$ by Lemma K.4. The first statement of the lemma follows because $C'(\cdot, i) > 0$ on $\mathcal{U}$. The second statement follows from the first and part (d) of Lemma K.2. □

For the next lemma, recall the multipliers $\eta_i(\mathbf{v})$ and $\sigma_{ij}(\mathbf{v})$ from the “interim” problem defined in SA-C.3.3 (just below [L $_i$]).

Lemma K.6. Suppose the environment is [TVC]-regular and $d = 2$. We have the following:

(a) $\mathbf{v} \in E_1$ if and only if

$$[\text{K.6}] \quad \eta_2(v_2) = \eta_1(\mathbf{v}) + \frac{\sigma_{21}(\mathbf{v})}{f_{12}}$$

(b) Fix $\mathbf{v} \in V_2$. At the optimum, we have

$$[\text{K.7}] \quad \eta_2(w_{12}(\mathbf{v}, s)) = \eta_1(\mathbf{w}_1(\mathbf{v}, s)) + \frac{\sigma_{21}(\mathbf{w}_1(\mathbf{v}, s))}{f_{12}} - \frac{(f_{11} - f_{21})}{f_{11}} \frac{\sigma_{21}(\mathbf{v})}{f_{12}}$$

(c) $\mathbf{v} \in E_2$ if and only if

$$[\text{K.8}] \quad \eta_2(v_2) = \eta_1(\mathbf{v}) + \frac{\sigma_{21}(\mathbf{v})}{f_{22}}$$

Proof. In the $d = 2$ case, [C.10] at $i = 2$ reduces to

$$[\text{K.9}] \quad \frac{\lambda_2(\mathbf{v}, s)}{f_{s2}} = \eta_2(\mathbf{v}) - \frac{f_{s1}}{f_{s2}} \sigma_{21}(\mathbf{v})$$

and [C.10] at $i = 1$ reduces to

$$[\text{K.10}] \quad \frac{\lambda_1(\mathbf{v}, s)}{f_{s1}} = \eta_1(\mathbf{v}) + \sigma_{21}(\mathbf{v})$$

Thus, by the envelope formulas [Env_{*i*}] and the definition of [$\tilde{\mathbf{E}}_i$], $\mathbf{v} \in E_s$ if and only if

$$\eta_2(\mathbf{v}) - \frac{f_{s1}}{f_{s2}} \sigma_{21}(\mathbf{v}) = \eta_1(\mathbf{v}) + \sigma_{21}(\mathbf{v})$$

Part (a) follows from rearranging the above display when $s = 1$. Part (c) follows from the same, when $s = 2$.

For part (b), first observe that a simple rearrangement of the optimality conditions [FOC_{*w_{ij}*}] for $i = 1$ delivers

$$[\text{K.11}] \quad \frac{P_1(\mathbf{w}_1(\mathbf{v}, s), 1)}{f_{11}} - \frac{P_2(\mathbf{w}_1(\mathbf{v}, s), 1)}{f_{12}} = \frac{\mu_{21}(\mathbf{v}, s)}{f_{s1}} \cdot \frac{f_{11} - f_{21}}{f_{11} \cdot f_{12}}$$

Subtracting [K.9] from [K.10], invoking the envelope formulas [Env_{*i*}], and combining with the above display yields

$$[\text{K.12}] \quad \eta_1(\mathbf{v}) + \sigma_{21}(\mathbf{v}) - \eta_2(\mathbf{v}) + \frac{f_{s1}}{f_{s2}} \sigma_{21}(\mathbf{v}) = \frac{\mu_{21}(\mathbf{v}, s)}{f_{s1}} \cdot \frac{f_{11} - f_{21}}{f_{11} \cdot f_{12}}$$

Now, in the $d = 2$ case, [C.11] reduces to $\sigma_{21}(\mathbf{v}) = \mu(\mathbf{v}, s)/f_{s1}$. (Thus, in particular, the weighted multiplier $\mu(\mathbf{v}, s)/f_{s1}$ is independent of s .) Plugging this in to the above display and rearranging yields part (b) of the lemma, completing the proof. $\square$

K.3. Proofs of Facts 1–4

We now prove each of the facts in turn.

Proof of Fact 1. This is an immediate corollary of part (b) of Lemma K.1. $\square$

We note that the properties in Fact 2 could be proved easily for any $d \geq 2$ by appealing to arguments in Thomas and Worrall (1990). We give a direct proof in the $d = 2$ case.

Proof of Fact 2. Part (a) follows from part (d) of Lemma K.2. Part (b) follows from Remark C.7. Consider part (c). Because D is a convex cone in $\mathcal{U}^d$ (part (a) of Theorem 3) and E is a ray in D (part (d) of Lemma K.2), it is easy to see that the ray E must be “upward sloping.” That is, for any $\mathbf{v}, \mathbf{v}' \in E$ such that $\mathbf{v} \neq \mathbf{v}'$, it must be that either $\mathbf{v} \ll \mathbf{v}'$ or $\mathbf{v}' \ll \mathbf{v}$. By Lemmas C.9 and K.4 (noting that $E_2 = E$ by part (d) of Lemma K.2), we have $D_1P(\mathbf{w}_2(\mathbf{v}, s), 2) > D_1P(\mathbf{v}, s)$. It then follows from part (b) of Lemma K.2, the fact that E is “upward sloping,” and [C.3] together with Lemma C.6(c) that $\mathbf{w}_2(\mathbf{v}, s) \gg \mathbf{v}$, as desired. Moreover, by the martingale property (Proposition 4.4), we have $D_1P(\mathbf{w}_1(\mathbf{v}, s), 1) < D_1P(\mathbf{v}, s)$. Remark C.7, together with the argument just used, then implies that $\mathbf{w}_1(\mathbf{v}, s) \ll \mathbf{v}$. This completes the proof of part (c). $\square$

Proof of Fact 3. Part (b) is a corollary of Lemma C.8. Consider part (c). By Lemmas C.9 and K.4, we have $D_1P(\mathbf{w}_2(\mathbf{v}, s), 2) > D_1P(\mathbf{v}, s)$. It then follows from part (b) of Lemma K.2, the fact that E_2 is “upward sloping,” and [C.3] together with Lemma C.6(c) that $\mathbf{w}_2(\mathbf{v}, s) \gg \mathbf{v}$, as desired. Finally, consider parts (a) and (d). For part (a), let $\mathbf{v} \in E_2$. Part (c) of Lemma K.6, Lemma K.4, and $f_{22} > f_{12}$ imply that

$$\eta_2(v_2) < \eta_1(\mathbf{v}) + \frac{\sigma_{21}(\mathbf{v})}{f_{12}}$$

But, by [K.9] and [K.10] (together with the envelope formulas [Env_{*i*}]), this is equivalent to $P_1(\mathbf{v}, 1)/f_{11} > P_2(\mathbf{v}, 1)/f_{12}$. This implies that E_1 lies strictly above E_2 because strict convexity and continuous differentiability of P (Proposition 3.2) imply that the lower contour sets of $P(\cdot, 1)$ are strictly convex, and tangent to the vector $\mathbf{f}_1^\perp := (1/f_{11}, -1/f_{12})$ exactly at points $\mathbf{v} \in E_1$. This completes the proof of part (a). The proof of part (d) is similar, invoking part (b) of Lemma K.6. $\square$

Proof of Fact 4. This is an immediate corollary of Lemma K.3. $\square$

References

Abreu, D., D. Pearce, and E. Stacchetti (1990). “Toward a Theory of Discounted Repeated Games with Imperfect Monitoring”. *Econometrica* 58.5, pp. 1041–1063.

- Bloedel, A. W., R. V. Krishna, and O. Leukhina (2025). “Insurance and Inequality with Persistent Private Information”. *Econometrica* 93.3, pp. 821–857.
- Echenique, F. (2005). “A Short and Constructive Proof of Tarski’s Fixed-Point Theorem”. *International Journal of Game Theory* 33.2, pp. 215–218.
- Gärtner, B. and J. Matoušek (2012). *Approximation Algorithms and Semidefinite Programming*. Springer-Verlag.
- Kamihigashi, T. (2014). “Elementary results on solutions to the Bellman equation of dynamic programming: existence, uniqueness, and convergence”. *Economic Theory* 56.2, pp. 251–273.
- Kreps, D. M. (1977). “Decision Problems with Expected Utility Criteria, I: Upper and Lower Convergent Utility”. *Mathematics of Operations Research* 2.1, pp. 45–53.
- Matoušek, J. and B. Gärtner (2007). *Understanding and Using Linear Programming*. New York: Springer.
- Stokey, N. L., R. E. Lucas Jr, and E. C. Prescott (1989). *Recursive Methods in Economic Dynamics*. Cambridge, Ma: Harvard University Press.
- Thomas, J. P. and T. Worrall (1990). “Income Fluctuation and Asymmetric Information: An Example of a Repeated Principal-Agent Problem”. *Journal of Economic Theory* 51, pp. 367–390.
- Vohra, R. V. (2005). *Advanced Mathematical Economics*. New York, NY: Routledge.