Supplementary Appendix to

Insurance and Inequality with Persistent Private Information

Alexander W. Bloedel∗ R. Vijay Krishna† Oksana Leukhina‡

7th September 2018

In this Supplementary Appendix to Bloedel, Krishna and Leukhina (2018), we provide proofs and constructions omitted from the main paper.

In Section S.1, we discuss the relation between sequential and recursive contracts, as well as conditions for their equivalence. Section S.2 constructs the optimal *full information* contract, while properties of the value function P are derived in Section S.3. Section S.4 constructs a suboptimal contract, and Section S.5 records some useful facts about pathwise properties of Markov chains.

S.1. Sequential and Recursive Contracts, and their Equivalence

S.1.1. Sequential Contracts

Here, we fill in the details concerning sequential contracts that are missing from Section 2.2. We remind the reader of the timing convention described in footnote [25] in the main text, whereby $\omega^{(t)} = \omega_i$ is equivalent to $s^{(t+1)} = i$, so that $s^{(t+1)}$ is the period t (not $t+1$!) type. While this convention is slightly awkward in the present sequential formulation, it is most natural in the recursive formulation adopted in the main paper; thus, for notational consistency, we maintain it here.

By the Revelation Principle, it is without loss to consider direct revelation mechanisms, whereby the principal chooses allocations conditional on histories of reported types. Let $\mathscr{G}$ denote the space of *private* histories, which are sequences of *realized* endowment types of the form $g = (s^1, \dots) \in S^\infty$.[1] Similarly, $\mathscr{H}$ is the space of *public* histories, which are sequences of

(∗) Stanford University. Email: abloedel@stanford.edu
(†) Florida State University. Email: rvk3570@gmail.com
(‡) Federal Reserve Bank of St. Louis. Email: oksana.m.leukhina@gmail.com
(1) $s^{(t)}$ is a random variable and s^t is a realization.

reported endowment types $h = (\hat{s}^0, \hat{s}^1, \dots) \in S^\infty$. (This is the same notation used in the main text to describe the space of paths. See Section 3.2.) Let G^t and H^t denote the spaces of length-t private and public histories, respectively. Thus, $h^t \in H^t$ is of the form $h^t = (s^1, \dots, s^t)$, so that h^t records the type realizations in periods $0, 1, \dots, t-1$. A *(pure) reporting strategy* for the agent, denoted by $\sigma := (\sigma_t)_{t=0}^\infty$ is a sequence of functions $\sigma_t : G^{t+1} \times H^t \to S$, which map past realized types, the current realized type, and past reports into a current reported type. The *truthful strategy* σ^* is defined by $\sigma_t^*\left((g^t, s^t), h^{t)}\right) = s^{t+1}$ for all $g^t \in G^t$ and $h^t \in H^t$. That is, it specifies truthfully reporting the current type regardless of the history of true and reported types.[2] We say that a strategy σ is *admissible* if $\sigma_t\left((g^t, s^{t+1}), h^t\right) \le s^{t+1}$ for all $g^t \in G^t$ and $h^t \in H^t$. That is, a strategy is admissible if the agent never over-reports his current endowment. The set of admissible strategies is denoted Σ. Under No Hidden Borrowing (Assumption NHB in the main text), the agent only has access to $\sigma \in \Sigma$.

While sequential contracts are naturally described in terms of transfers of the consumption good, it is most convenient to formulate them in terms of flow utilities. A transfer of c_i from the principal to an agent with endowment ω_i delivers to the agent flow utility $u_i := U(c_i + \omega_i)$. Thus, any such transfer is equivalent to a flow utility allocation of u_i to an i-type agent at cost $C(u_i, i) := C(u_i) - \omega_i$, where $C(u) := U^{-1}(u)$. Thus, a *sequential contract*, denoted $\tilde{u} := \left(u^{(t)}\right)_{t=0}^\infty$, is a $\mathcal{U}$-valued stochastic process adapted to the filtration implied by the public histories and the initial vector of promises $\mathbf{v}^(0)$. In particular, given some $\mathbf{v}^{(0)}$, $u^{(t)}$ is H^{t+1}-measurable — ie, it depends on all past and current reports. Let $u_i^{(t)}$ denote the (random) date t flow allocation when the period t report $\hat{s}^{(t+1)} = i$. The set of all sequential contracts is $\mathcal{A}$. Recall that $\psi : \mathcal{U} \times S \times S \to \mathcal{U}$ defined as $\psi(u, i, j) := U(\omega_i + C(u, j))$ specifies how an agent of type i values the flow utility allocation intended for type j. In particular, if an agent of type i lies and claims to be of type $j \neq i$, he receives flow utility $\psi(u_j, i, j)$. If he truthfully reports his type to be i, he receives flow utility $\psi(u_i, i, i) = u_i$.

Every sequential contract $\tilde{u}$ and reporting strategy σ together induce a stochastic process $\left(s^{(t+1)}, \hat{s}^{(t+1)}, u^{(t)}\right)_{t=0}^\infty$ over tuples of true types, reported types, and flow utility allocations in each period. Taking $\tilde{u}$ as given, denote the law of this process by $\mathbf{P}^\sigma \in \Delta\left(S^\infty \times S^\infty \times \mathcal{U}^\infty\right)$ and its associated expectation operator by $\mathbf{E}^\sigma[\cdot]$. Thus, the agent's preferences over sequential contracts and admissible reporting strategies are represented by the lifetime utility function $\hat{U} : \mathcal{A} \times \Sigma \times S \to \mathcal{U} \cup \{-\infty\}$ defined by

$$\hat{U}\left(\tilde{u}, \sigma, s\right) := \mathbf{E}^\sigma\left[\sum_{t=0}^\infty \alpha^t \psi\left(u^{(t)}_{\hat{s}^{(t+1)}}, s^{(t+1)}, \hat{s}^{(t+1)}\right) \middle| s^{(0)} = s\right]$$

When the agent follows the strategy σ and the principal has prior belief $\mu \in \Delta(S)$ over the

(2) In our Markovian environment, this is no more restrictive than requiring that the agent report truthfully conditional on having done so in the past. See, eg, Section 3.3 of Pavan, Segal and Toikka (2014) for a discussion of this point.

agent's initial type, the cost to the principal of a sequential contract $\tilde{u}$ is

$$R(\tilde{u}, \sigma, \mu) := \mathbf{E}^{\sigma}_{s^{(0)} \sim \mu}\left[\sum_{t=0}^{\infty} \alpha^t C(u^{(t)}_{\hat{s}^{(t+1)}}, \hat{s}^{(t+1)})\right]$$

(We will assume for simplicity that costs are always well-defined.)

We say that a sequential contract $\tilde{u}$ *implements* $\mathbf{v} \in \mathcal{U}^d$ if it satisfies

[S-PK$_i$]
$$v_i = \hat{U}(\tilde{u}, \sigma^*, i)$$

for all $i \in S$ and

[S-IC]
$$\hat{U}(\tilde{u}, \sigma^*, i) \geq \hat{U}(\tilde{u}, \sigma, i) \qquad \forall\, \sigma \in \Sigma$$

The collection of [**S-PK$_i$**] constraints are the familiar *promise-keeping* conditions, which ensure that the principal actually delivers the appropriate level of contingent lifetime utility to the agent. The *incentive compatibility* constraint [**S-IC**] requires that truthtelling is an optimal reporting strategy for the agent.

The set of sequential contracts that implement $\mathbf{v}$ is $\Pi(\mathbf{v})$. The principal's objective is to minimize the expected, discounted cost of delivering $\mathbf{v}$ to the agent. Her value function is therefore

[SP]
$$P^*(\mathbf{v}, \mu) := \inf_{\tilde{u} \in \Pi(\mathbf{v})} R(\tilde{u}, \sigma^*, \mu)$$

We refer to this as the principal's *sequence problem*. (Note that the principal's costs are evaluated under the truthtelling measure.) A sequential contract is *sequentially optimal* if it attains the infimum in [**SP**].

To connect this to the usual efficiency problem, we must allow the principal to *initialize* the initial vector $\mathbf{v}$ of contingent promises subject to a constraint on *expected* promised utility. Formally, the principal's *Pareto problem* is given by

[Pareto]
$$\begin{aligned} K^*(w, \mu) := \inf_{\mathbf{v} \in \mathcal{U}^d} & \; P^*(\mathbf{v}, \mu) \\ \text{s.t.} \qquad & \mathbf{E}^{\mu}[\mathbf{v}] \geq w \end{aligned}$$

Thus, by varying $w \in \mathcal{U}$, we trace out the entire Pareto frontier. [**Pareto**] is the counterpart for [**SP**] of the efficiency problem [**Eff$_i$**], which was defined in Section 5.1 for the recursive problem [**RP**]. As is well known, this cost minimization problem is dual to the "primal" problem of maximizing a utilitarian welfare functional subject to an resource constraint (see, eg, Golosov, Tsyvinski and Werquin (2016)). We find it convenient to work directly with the dual, skipping the standard step of relating the two problems.

S.1.2. Proof of Lemma 3.2

Proof. The converse statement at the end of the lemma is obvious. Part (c) follows from parts (a) and (b). For part (a), let $\xi \in \Xi^*(\mathbf{v})$. Iterating forward T times on the recursive promise keeping constraints [**PK**$_i$] gives

$$v_i = \mathbf{E}\left[\sum_{t=0}^{T-1} \alpha^t \tilde{u}_\xi^{(t)} \mid s^{(0)} = i\right] + \mathbf{E}\left[\alpha^T \, \mathbf{E}^{\mathbf{f}_{s^{(T+1)}}}\left[\mathbf{v}_\xi^{(T+1)}\right] \mid s^{(0)} = i\right]$$

for all $i \in S$. The Monotone Convergence Theorem implies that

$$\lim_{T\to\infty} \mathbf{E}\left[\sum_{t=0}^{T} \alpha^t \tilde{u}_\xi^{(t)} \mid s^{(0)} = s_j\right] = \mathbf{E}\left[\sum_{t=0}^{\infty} \alpha^t \tilde{u}_\xi^{(t)} \mid s^{(0)} = s_j\right]$$

and the Bounded Convergence Theorem (which applies under [**TVC**]) implies that

$$\lim_{T\to\infty} \mathbf{E}\left[\alpha^T \, \mathbf{E}^{\mathbf{f}_{s^{(T+1)}}}\left[\mathbf{v}_\xi^{(T+1)}\right] \mid s^{(0)} = i\right] = 0$$

Combining the above displays delivers part (a). Now, part (b) is essentially Lemma 2 in Green (1987). The proof of part (b) standard and nearly identical to the one in Green (1987), and thus omitted. □

S.1.3. Equivalence of Sequential and Recursive Problems

We have introduced two versions of the principal's problem: the sequential problem [**SP**] from Appendix S.1.1 and the recursive problem [**RP**] from Section 3.1. Here, we describe the relation between these two formulations and our reasons for focusing on [**RP**], elaborating on the informal discussion in Section 3.2.

It is well known that every recursive contract generates a unique sequential contract (namely, the induced allocation) and that every "stationary" sequential contract induces a unique recursive contract. Moreover, it is without loss of optimality in [**SP**] to restrict attention to "stationary" sequential contracts.[3]

The difference lies in the definition of the feasible sets. It is easy to see that the recursive contract induced by a "stationary" sequential contract satisfies the recursive constraints at each step. The converse is not necessarily true. Given a recursive contract $\xi \in \Xi(\mathbf{v})$, the induced allocation $\tilde{u}_\xi$ need not be an element of $\Pi(\mathbf{v})$ for two reasons. First, if the promised utility

(3) Every sequential contract induces a stochastic process for contingent promised utilities through conditional expectations of future flow utility allocations. Informally, a sequential contract is "stationary" if the continuation allocation depends on the history only through the implied vector of contingent promised utilities starting at that history. That "stationarity" is without loss of optimality can be deduced from arguments analogous to those used in the proof of Lemma A.2 in Atkeson and Lucas (1992).

process grows too quickly and violates the condition [**DP**] from Lemma 3.2 — essentially, if the principal violates the analogue of a no Ponzi scheme condition — then $\tilde{u}_\xi$ may fail to satisfy the sequential promise keeping constraints [**S-PK**$_i$]. That is, it may fail to *deliver promises*. Second, even if $\tilde{u}_\xi$ delivers promises, it may not satisfy the sequential incentive compatibility constraint [**S-IC**]. This is because recursive incentive compatibility only deters one-shot deviations, and without the one-stage deviation principle it is possible that strategies involving infinitely-many deviations may strictly dominate truthtelling.

Under certain conditions, these problems do not arise, in which case [**SP**] and [**RP**] are equivalent. If the agent's utility function is bounded, then it is obvious that promises are delivered and that the one-stage deviation principle applies (see, eg, Theorem 2.1 in Fernandes and Phelan (2000) for the standard argument). Even when the agent's utility function is unbounded, one can show that the two problems are equivalent by establishing, for instance, that (i) both value functions — namely, P^* and P — satisfy the same Bellman equation, (ii) this Bellman equation has a unique solution within some class of functions, and (iii) both P and P^* lie in that class. Then, it follows that the value functions coincide and thus generate the same policy functions; either the policy functions generate a solution to [**SP**] or they do not, in which case [**SP**] does not have a solution.

Unfortunately, neither of these approaches appear to work in our setting. The agent's utility function is unbounded below, and standard contraction mapping arguments do not seem to apply to the principal's Bellman equation, leaving open the possibility that, while P^* and P are both solutions, they do not coincide.[4] We view this as a purely technical gap, as it arises from the multi-dimensionality of the state space and has little to do with the economics at hand. Nevertheless, it is important to understand how this gap affects the analysis.

This leaves us with two options. First, we can show that P^* satisfies the Bellman equation [**FE**] and characterize the recursive contract induced by the policy functions.[5] This approach is useful because the sequential problem [**SP**] is the "correct" one from an economic standpoint, in that it incorporates the full set of constraints. If the policy functions induce a sequential

(4) Thomas and Worrall (1990) encounter similar issues and show how to adapt contraction arguments to their setting. Their idea is to construct two functions that bound the value function — one from below, the other from above — and show that the difference between the upper and lower bound is itself bounded. Then, when searching for the value function, one may restrict attention to functions that lie (pointwise) between the upper and lower bound, and standard arguments from bounded dynamic programming apply. As we have seen in part (b) of Theorem 2, the recursive value function P is unbounded even on bounded sets. While we can always construct a lower bound (we use the first-best value function Q^* described in Appendix S.2, though one could also consider $T^n Q^*$, the n-fold iterate of the Bellman operator [**T**] on Q^*) and can construct an upper bound in at least one special case (part (f) of Theorem 1 or Proposition S.4.7), we cannot guarantee that the difference between them is bounded. Thus, it seems unlikely that similar techniques apply to our setting, at least without detailed study of growth-rate properties of these upper and lower bounding functions.

(5) This is true, and follows from arguments similar to those used in the proof of Theorem 1 in Farhi and Werning (2007).

contract in $\Pi(\mathbf{v})$, then we will have characterized the solution to [**SP**]; if not, then [**SP**] does not have a solution in the first place. But verifying this last piece involves checking that the induced sequential contract satisfies a transversality-type condition, and this is typically not possible to do analytically.[6] Moreover, it is sometimes more difficult to establish important properties when using P^* instead of P. For example, the proof that the directional derivative $\mathsf{D}_1 P(\cdot, s) \geq 0$ (see Lemma B.1) is based on the fact that P is the *smallest* solution to [**FE**] in an appropriate space (part (a) of Theorem 2). In contrast, P^* need not be the smallest (in an appropriate sense) solution to [**FE**], and it is not obvious how to establish that $\mathsf{D}_1 P^*(\cdot, s) \geq 0$.

Second, we can study solutions ξ^* to the recursive problem [**RP**] directly and give conditions under which the indued allocation $\tilde{u}_{\xi^*} \in \Pi(\mathbf{v})$. Because, as described above, [**RP**] has fewer constraints than [**SP**] and is thus a (weak) relaxation of [**SP**]. Thus, under such conditions on ξ^*, $\tilde{u}_{\xi^*}$ solves [**SP**] and, from the standpoint of optimality, the two problems are equivalent.[7] We adopt this second approach for three reasons. First, the sufficient condition for $\tilde{u}_{\xi^*} \in \Pi(\mathbf{v})$ — namely, [**TVC**]-implementability, see Lemma 3.2 — is no harder or easier to verify than the transversality-type condition one obtains when working with P^*. Second, the value function P from [**RP**] is the function that can be studied numerically by standard iterative procedures, by virtue of being the smallest solution to [**FE**].[8] Third, once one adopts the recursive perspective, [**RP**] is arguably the "correct" problem to study from a mathematical and conceptual standpoint, as it is formulated directly in terms of recursive objects.

S.2. Optimal Full-Information Contract

We now characterize the unique full information optimal (recursive) contract. First, we define the feasible set for the full information problem. Namely, for each $\mathbf{v} \in \mathcal{U}^d$, we have

[**S.2.1**] $$\Gamma^{\text{FB}}(\mathbf{v}) := \left\{(u_i, \mathbf{w}_i)_{i \in S} \in (\mathcal{U} \times \mathcal{U}^d)^d : [\textbf{PK}_i] \text{ holds for all } i \in S\right\}$$

(6) This is the approach used in, eg, Farhi and Werning (2007). They give sufficient conditions under which the policy functions solve the sequential problem, but their proofs rely heavily on special properties of their value function. These special properties are not satisfied in the setting of Thomas and Worrall (1990), who carry out the verification step only in the special case of CARA utility, where (near) closed forms are available. Even with CARA utility, the multi-dimensionality of our problem prevents us from obtaining such an explicit description of the optimal contract.

(7) This is the approach taken by Green (1987) and, with minor modifications, Atkeson and Lucas (1992). More generally, it is in the spirit of "verification theorems" used throughout stochastic control.

(8) In particular, starting from the first-best value function Q^*, one numerically computes the iterates $T^n Q^*$ (recall the Bellman operator T defined in [**T**]) using value function iteration. *If* $T^n Q^*$ converges to some function Q^∞, and *if* Q^∞ is a fixed point of T, then $Q^\infty = P$. (See part (a) of Theorem 2 and Theorem 4.14 in Stokey, Lucas and Prescott (1989, pp. 92-93).) While we cannot guarantee that Q^∞ is a fixed point of T in general, this can be (approximately) verified numerically.

$$\text{[FB]} \qquad Q^*(\mathbf{v}, s) := \inf_{\xi \in \Xi^{FB}(\mathbf{v})} \mathsf{E}\left[\sum_{t=0}^{\infty} \alpha^t C\left(u_\xi^{(t)}, s^{(t+1)}\right) \Big| s^{(-1)} = s\right]$$

We begin with an important technical lemma.

Lemma S.2.1. The first-best value function $Q^* : \mathcal{U}^d \times S \to \mathbb{R}$ satisfies the functional equation

$$\text{[S.2.2]} \qquad Q^*(\mathbf{v}, s) = \inf_{(u_i, \mathbf{w}_i)_{i \in S} \in \Gamma^{FB}(\mathbf{v})} \sum_{i \in S} f_{si} \left[C(u_i, i) + \alpha Q^*(\mathbf{w}_i, i)\right]$$

and each $Q^*(\cdot, s)$ is convex and continuously differentiable. Moreover, the infimum in [**S.2.2**] is attained at each $(\mathbf{v}, s) \in \mathcal{U}^d \times S$.

Proof. That Q^* satisfies [**S.2.2**] is standard, so the proof is omitted. It is also immediate that Q^* is finite. Proofs of the other properties are standard and are laid out for the (private information) valued function P in Supplementary Appendix S.3. See, in particular, Lemma S.3.7 for convexity, Lemmas S.3.8 and S.3.9 for attainment of the infimum, and Lemma S.3.11 for continuous differentiability. □

With Lemma S.2.1 in hand, solving for the optimal policy in [**S.2.2**] reduces to a smooth, convex, finite-dimensional minimization problem. It is easy to see that it is never optimal to set $u_i = 0$ because of the Inada condition $C(0, i) = \infty$. Therefore, we may substitute u_i out of the promise keeping constraint [**PK**$_i$] and instead optimize over all $\mathbf{w}_i \in \mathcal{U}^d$ for each i in [**S.2.2**].

With this substitution, the first order condition for w_{ij} is

$$\text{[FB-FOC}w_{ij}\text{]} \qquad Q_j^*(\mathbf{w}_i, i) = f_{ij} \cdot C'(u_i, i)$$

for all $i, j \in S$, and the envelope condition is

$$\text{[FB-Env}_j\text{]} \qquad Q_j^*(\mathbf{v}, s) = f_{sj} \cdot C'(u_j, j)$$

for all $j \in S$. From these optimality conditions, it is easy to deduce the following characterization of the full-information optimum.

Lemma S.2.2. Fix any initial promise $\mathbf{v} \in \mathcal{U}^d$. There exists a unique recursively optimal recursive contract. If the first type is $s^{(0)} = s \in S$, the optimal contract satisfies

$$\text{[S.2.3]} \qquad u_i^{(t)}(\mathbf{v}, s^{(0)} = s) = (1 - \alpha) \cdot v_s \qquad \text{for all } i \in S$$

$$\text{[S.2.4]} \qquad \mathbf{w}_i^{(t)}(\mathbf{v}, s^{(0)} = s) = v_s \mathbf{1} \qquad \text{for all } i \in S$$

The value function $Q^*(\cdot, s)$ is strictly convex for each $s \in S$, the derivative is strictly positive $\mathsf{D}Q^*(\mathbf{v}, s) \gg \mathbf{0}$, and thus $\mathsf{D}_1 Q^*(\mathbf{v}, s) > 0$ for all $(\mathbf{v}, s) \in \mathcal{U}^d \times S$.

The optimal policy follows easily from the first order and envelope conditions above and is omitted. That $Q^*(\cdot, s)$ is strictly convex also follows from the optimal policy. Finally, the positive partial derivatives are immediate from the envelope condition [**FB-Env**$_j$].

Consider also the *full-information efficiency problem*

$$\text{[Eff}_i^{\text{FB}}\text{]} \qquad \begin{aligned} K^*(v,s) := \min_{\mathbf{v}\in\mathcal{U}^d} \; & Q^*(\mathbf{v},s) \\ \text{s.t.} \quad & \mathsf{E}^{\mathbf{f}_s}[\mathbf{v}] \geq v \end{aligned}$$

for each $i \in S$. This is the full-information analogue of the efficiency problem [**Eff**$_i$] defined in Section 5.1. The following characterization follows immediately from the preceding lemmas and the primitive model assumptions stated in Section 2.1 and 3.1.

Lemma S.2.3. For each $s \in S$ and $v \in \mathcal{U}$, the full-information efficiency problem [**Eff**$_i^{\text{FB}}$] has the unique solution

$$\text{[S.2.5]} \qquad \mathbf{v}^*(v,s) = v \cdot \mathbf{1}$$

which is continuous in $v \in \mathcal{U}$. Thus, the value function for this problem is given by[9]

$$\text{[S.2.6]} \qquad K^*(v,s) = C((1-\alpha)v) - \mathbb{E}\left[\sum_{t=0}^{\infty} \alpha^t \omega^{(t)} \,\middle|\, s^{(0)} = s\right]$$

which is clearly strictly increasing, strictly convex, and continuously differentiable, and satisfies the Inada conditions $\lim_{v\to-\infty} K^{*\prime}(v,s) = 0$ and $\lim_{v\to 0} K^{*\prime}(v,s) = +\infty$.

S.3. Proofs for Sections 4.3 and 4.4

In this appendix, we collect all proofs related to the regularity conditions from Sections 4.3 and the Bellman equation from Section 4.4. In particular, we first prove Lemma 4.2 in Appendix S.3.1. Then, in Appendix S.3.2 we prove Theorem 2. Half of part (c) of Theorem 2 is proved separately in Appendix B.3.3.

S.3.1. Proof of Lemma 4.2

We begin with the proof of part (a) of Lemma 4.2. (Note that, if $\mathbf{v} \in V_d$, then part (b) of Lemma A.5 and Lemma A.6 together imply that $\Gamma_\circ(\mathbf{v}) \neq \emptyset$. The following proof extends the same reasoning to the potential case in which $D \supsetneq V_d$.)

Proof of Part (a), Case (i). Let the Markov 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.

(9) Recall from Section 3.1 that $C(u) := U^{-1}(u)$ while $C(u,i) := C(u) - \omega_i$.

Let $\ell \in S$ be given. By hypothesis, $(u_\ell, \mathbf{w}_\ell) \in \mathcal{U} \times D$ satisfies [$\mathbf{PK}_i$] $(i = \ell)$ and each of the [$\mathbf{IC}_{ij}$] $(i > j = \ell)$. Let $\varepsilon > 0$, and let $w'_{\ell k} := w_{\ell k} - \varepsilon/2^k$ for $k = 1, \ldots, d$. Now, let $u'_\ell := u_\ell + \delta$ such that $\delta = \alpha\varepsilon\, \mathsf{E}^{\mathbf{f}_\ell}[(\frac{1}{2}, \ldots, \frac{1}{2^d})]$. By construction, $(u'_\ell, \mathbf{w}'_\ell)$ satisfies [$\mathbf{PK}_i$] $(i = \ell)$. 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 1).

Recall from Lemma A.1 that for all $i > \ell$ we have $\psi'(u, i, \ell) < 1$ since $u < 0$, which means that $\psi(u_\ell + \delta, i, \ell) - \psi(u_\ell, i, \ell) = \int_0^\delta \psi'(u + y, i, \ell)\, \mathrm{d}y < \delta$. Therefore, for all $i > \ell$, we have

$$\text{[S.3.1]} \qquad \left[\psi(u_\ell + \delta, i, \ell) - \psi(u_\ell, i, \ell)\right] - \alpha\varepsilon\, \mathsf{E}^{\mathbf{f}_i}[(\tfrac{1}{2}, \ldots, \tfrac{1}{2^d})] < 0$$

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

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

where the weak inequality is because $(u_\ell, \mathbf{w}_\ell)$ satisfies [$\mathbf{IC}_{ij}$] $(i > j = \ell)$, and the strict inequality is because of [S.3.1]. Thus, it is clear that $(u'_\ell, \mathbf{w}'_\ell) \in \Gamma_\circ(\mathbf{v})$. □

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 $\mathrm{co}\left(\{\mathbf{f}_i\}_{i=1}^d\right)$. Thus, by the Separating Hyperplane Theorem, for each $\ell \in S$ there exists a vector $\mathbf{a}_\ell \in \mathbb{R}^d$ such that $\mathsf{E}^{\mathbf{f}_i}[\mathbf{a}_\ell] > \mathsf{E}^{\mathbf{f}_\ell}[\mathbf{a}_\ell] > 0$ for all $i > \ell$. We may then adopt exactly the same proof used in Case (i) above by simply replacing the vector $(\frac{1}{2}, \ldots, \frac{1}{2^d})$ with the vectors $\mathbf{a}_\ell$ when perturbing $(u_\ell, \mathbf{w}_\ell)$. □

Towards the proof of part (b) of Lemma 4.2, we next present a series of lemmas. Lemma S.3.1, stated below, does not rely on the CARA or PPR/MLRP assumptions.

Lemma S.3.1. There exists a real-valued function $Q^L : D \times S \to \mathbb{R}$ such that $Q^L \le P$ on $D \times S$.

Proof. Lemma S.2.1 states that the first-best value function Q^* is real-valued on $\mathcal{U}^d \times S$. Let Q^L denote the restriction of Q^* to $D \times S$. By definition of $Q*$ and P, we have that $Q^L \le P$ on $D \times S$. □

Lemma S.3.2. Suppose that the agent has CARA utility and that the Markov process for endowments satisfies either MLRP or PPR. Then there exists a real-valued function $Q^U : D \times S \to \mathbb{R}$ such that $P \le Q^U$ on $D \times S$.

Proof. By part (f) of Theorem 1, 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 Appendix S.4. By Proposition S.4.7, the value function induced by the contract ζ is real-valued, $Q^\zeta : D \times S \to \mathbb{R}$. Moreover, by the definition of Q^ζ and P, we have that $P \leq Q^\zeta$ on $D \times S$. Thus, let $Q^U := Q^\zeta$. □

Lemma S.3.3, stated below, does not rely on the CARA or PPR/MLRP assumptions. It will also be used later in the proof of Theorem 2.

Lemma S.3.3. There exists some $\eta \in \mathbb{R}_{++}$ such that, for each $\mathbf{v} \in D$ and $(u_i, \mathbf{w}_i)_{i \in S} \in \Gamma(\mathbf{v})$,

$$u_i \in (v_i, 0)$$
$$w_{ij} \in (\eta v_i, 0)$$

for all $i, j \in S$.

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

Lemma S.3.4. 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, $U(c) = -e^{-c}$ (we set the risk aversion coefficient $\rho := 1$ for notational simplicity), the cost function is given by $C(u) := U^{-1}(u) = -\log(-u)$. Let $\theta_i = e^{-\omega_i}$. Then it is easy to see that $C(u, i) = C(u/\theta_i)$. Fix an initial promise $\mathbf{v} \in \mathscr{U}^d$, a seed type $s \in S$, and a recursive contract $\xi \in \Xi^{\mathrm{FB}}(\mathbf{v})$. 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 ξ initialized at $\mathbf{v}$. Fix a path $h \in \mathscr{H}$. Recall the constant $\eta > 0$ defined in Lemma S.3.3. Along the path h, we have

$$C(u_i^{(t)}(h)) = -\log\left(\left|\frac{v_i^{(t)}(h) - \alpha \mathbb{E}^{\mathbf{f}_i}\left[\mathbf{w}_i^{(t)}(h)\right]}{\theta_i}\right|\right)$$

$$\text{[S.3.2]} \qquad \geq -\log\left(\frac{1}{\theta_i} \cdot \left[|v_i^{(t)}(h)| + \sum_{j=1}^{d} |w_{ij}^{(t)}(h)|\right]\right)$$

$$\text{[S.3.3]} \qquad \geq \log(\min_{i \in S} \theta_i) - \log\left(\eta^t \cdot \max_{i \in S} |v_i^{(0)}| \cdot [1 + d \cdot \eta]\right)$$

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 S.3.3. 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 $h \in \mathscr{H}$ and the current type $i \in S$. Thus, taking the α-discounted sum, it is easy to see that there exist $\kappa_1, \kappa_2 \in \mathbb{R}$ such that

$$\text{[S.3.4]} \qquad \inf_{\xi \in \Xi^{\mathrm{FB}}(\mathbf{v}),\ h \in \mathscr{H}} \sum_{t=0}^{\infty} \alpha^t C\left(u_{s^{(t+1)}(h)}^{(t)}(h), s^{(t+1)}(h)\right) \geq \kappa_1 + \kappa_2 \log(\max_{i \in S} |v_i|)$$

The discounted sum on the LHS is the realized cost to the principal of ξ along path h. The inequality in [S.3.4] gives a lower bound on *pathwise* costs that is uniform across paths and feasible 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 [S.3.4] is that

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

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

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

Iterating on the same growth-rate bounds from Lemma S.3.3 used to derive [S.3.3], we may bound the RHS of the above inequality from below as follows:

$$\text{[S.3.5]} \qquad \kappa_1 + \kappa_2 \log(\max_i |v_i^{(t)}(h)|) \geq \kappa_1 + \kappa_2 \cdot t \cdot \log(\eta) + \kappa_2 \cdot \log(\max_{i \in S} |v_i^{(0)}|)$$

The RHS of [S.3.5] is independent of the path $h \in \mathscr{H}$. Then, using the fact that $\alpha^t \cdot t \to 0$, we have

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

which is the desired condition. □

We now collect these pieces into a proof of part (b) of Lemma 4.2.

Proof of Part (b). Lemmas S.3.1 and S.3.2 together imply that P must be real-valued whenever it is well-defined. Note that the inequality [S.3.4] in the proof of Lemma S.3.4 implies that P is everywhere well-defined. In particular, [S.3.4] implies that the lifetime discounted costs generated by any feasible recursive contract may be assumed non-negative along *each* path $h \in \mathscr{H}$ (up to the addition of a constant that depends only on the parameters κ_1, κ_2 and the initial promise $\mathbf{v} \in D$). Thus, the *expected* lifetime cost of any feasible recursive contract — and, thus, also P — is well-defined. This establishes that Condition R.1 is satisfied. Condition R.2 is satisfied by Lemma S.3.4. Condition R.3 is satisfied by case (i) of part (a) of Lemma 4.2, proved above. □

S.3.2. Proof of Theorem 2

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 $\overline{\mathbb{R}}^{V_d \times S}$ denote the space of functions $f : D \times S \to \overline{\mathbb{R}}$. Define the Bellman operator $T : \overline{\mathbb{R}}^{D \times S} \to \overline{\mathbb{R}}^{D \times S}$ by

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

Lemma S.3.5. Suppose that Condition R.1 holds.[10] Then the principal's value function, P, satisfies the Bellman equation

$$P = TP \qquad \text{[BE]}$$

Proof. This follows from standard Principle of Optimality arguments, eg, a straightforward adaptation of Theorem 4.2 of Stokey, Lucas and Prescott (1989, p. 71). □

Lemma S.3.6. Suppose that Conditions R.1 and R.2 hold. Then P is the smallest fixed point of the Bellman operator [T] in the order interval $[Q^*, \infty]$ (endowed with the pointwise order).

Proof. By Lemma S.3.5, 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 $\overline{v}$ 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. □

Lemma S.3.7. Suppose that Condition R.1 holds. Then $P(\cdot, s)$ is convex and continuous on D for each $s \in S$.

Proof. Part (b) of Theorem 1 states that the constraint correspondence $\Gamma(\cdot)$ has a convex graph. Thus, the feasibility correspondence $\Xi(\cdot)$ in [RP] has convex graph (when convex combinations of recursive contracts are defined pointwise). In addition, the return functions $C(\cdot, s)$ are convex. It follows from the definition of [RP] that $P(\cdot, s)$ is convex. Continuity of D follows immediately, as D is open by part (a) of Theorem 1. □

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

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

(10) The requirement of finiteness in Condition R.1 is not, strictly speaking, necessary. But it is sufficient and, in particular, rules out the possibility that P takes on both values $\pm\infty$, in which case the operator [T] need not be well-defined.

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

$$\Gamma^*(\mathbf{v}, s) := \arg \min_{(\boldsymbol{u},\mathbf{w})\in\Gamma(\mathbf{v})} H(P, (\boldsymbol{u}, \mathbf{w}), s) \tag{S.3.6}$$

Lemma S.3.8. Suppose that Condition R.1 holds. For any convergent sequence $(\mathbf{v}^n)_{n=0}^{\infty} \subset D$ with limit point $\mathbf{v}^* \in \operatorname{bd} D$, the boundary of D, we have $P(\mathbf{v}^n, s) \to +\infty$ for each $s \in S$.

Proof. Let $s \in S$ be fixed throughout. Let $(\boldsymbol{u}^n, \mathbf{w}^n) := (u_i^n, \mathbf{w}_i^n)_{i\in S}$ denote a sequence of menus with the following properties. First, $(\boldsymbol{u}^n, \mathbf{w}^n) \in \Gamma(\mathbf{v}^n)$ for all $n \in \mathbb{N}$. Now, because the sequence $(\mathbf{v}^n)$ converges and is thus bounded, it follows from Lemma S.3.3 that there exists a bounded set $R \subset (\mathcal{U} \times D)^d$ such that $(\boldsymbol{u}^n, \mathbf{w}^n) \in R$ for all $n \in \mathbb{N}$. Second, suppose that there exists some $\varepsilon > 0$ such that each $(\boldsymbol{u}^n, \mathbf{w}^n)$ is ε-optimal, ie,

$$H^n := H(P, (\boldsymbol{u}^n, \mathbf{w}^n), s) \leq P(\mathbf{v}, s) + \varepsilon \tag{S.3.7}$$

for all $n \in \mathbb{N}$. Such an $\varepsilon > 0$ and a corresponding sequence of menus exists by virtue of Lemma S.3.5.

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

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

To see this, observe that $\inf_{i\in S,\ (\boldsymbol{u},\mathbf{w})\in \operatorname{cl} R} Q^*(\mathbf{w}_i, i) > -\infty$ by Lemma S.2.1, 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 $\boldsymbol{u}^n \to \mathbf{0}$. This follows from the structure of the domain D — see, in particular, the statements and proofs of Lemma A.23 and Proposition A.24. It is then easy to see that the claim that$H^n \to +\infty$ follows, as we have

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

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

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 recursively 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 S.3.9. Suppose that Condition R.1 holds. 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

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

Define $K(\mathbf{v}, s) :=$ cl $L(\mathbf{v}, s)$, the closure of $L(\mathbf{v}, s)$. It follows from Lemma S.3.3 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 S.3.5), and thus there exist ε-optimal menus for each $\varepsilon > 0$. Moreover, it is easy to see that the function $H(P, \cdot) : \Gamma(\mathbf{v}) \to \mathbb{R}$ is convex, and thus that $K(\mathbf{v}, s)$ is a convex set (since $\Gamma(\mathbf{v})$ is convex by part (b) of Theorem 1).

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 $K(\mathbf{v}, s) \subseteq$ cl $\Gamma(\mathbf{v})$. If $K(\mathbf{v}, s) \cap (\text{cl } \Gamma(\mathbf{v}) \backslash \Gamma(\mathbf{v})) \neq \emptyset$, then there exists some sequence $((\boldsymbol{u}^n, \mathbf{w}^n))_{n=0}^{\infty} \subseteq L(\mathbf{v}, s)$ such that $(\boldsymbol{u}^n, \mathbf{w}^n) \to \text{bd } (\mathcal{U} \times D)^d$. (It is easy to see that $(\text{cl } \Gamma(\mathbf{v}) \backslash \Gamma(\mathbf{v})) \subseteq \text{bd } (\mathcal{U} \times D)^d$.) By the same arguments as in the proof of Lemma S.3.8, $H(P, (\boldsymbol{u}^n, \mathbf{w}^n)) \to +\infty$, 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 S.3.7), 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

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

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 [**S.3.6**]) 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. □

Lemma S.3.10. Suppose that Conditions R.1 and R.2 hold. Any conserving contract ξ^* is recursively optimal, ie, is optimal in the principal's recursive problem [**RP**]. Hence, there exists a recursively optimal contract.

Proof. Lemma S.3.9 shows that a conserving contract exists under Condition R.1, so it suffices to show that any such contract is recursively optimal for the principal. But recursively 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 S.3.5) and the infimum is attained at each step (Lemma S.3.9), induction gives

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

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

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

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 $t \geq T$,

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

Thus, we may apply Fatou's Lemma to get

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

Combining the above displays, we get

$$\text{[S.3.8]} \qquad P(\mathbf{v}, s) \geq \liminf_{T\to\infty} \mathbf{E}\left[\sum_{t=0}^{T} \alpha^t C(u^{(t)}, s^{(t+1)}) \Big|\, s^{(0)} = s\right]$$

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

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

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

and thus,

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

Then, by Fubini's Theorem,

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

Then it follows from [**S.3.8**] that $P(\mathbf{v}, s) \geq \mathbf{E}\left[\sum_{t=0}^{\infty} \alpha^t C(u^{(t)}, s^{(t+1)}) \,\Big|\, s^{(0)} = s\right]$. The opposite inequality holds by the definition of P in [**RP**]. Thus, $P(\mathbf{v}, s) = \mathbf{E}\left[\sum_{t=0}^{\infty} \alpha^t C(u^{(t)}, s^{(t+1)}) \,\Big|\, s^{(0)} = s\right]$ and hence the conserving contract is recursively optimal. □

Lemma S.3.11. Suppose that Conditions R.1 and R.2 hold. Then $P(\cdot, s)$ is continuously differentiable on D for each $s \in S$.

Proof. From Lemma S.3.9, the infimum in the Bellman equation is attained. Moreover, the domain D is open by part (a) of Theorem 1. The claim then follows from a straightforward adaptation of the Benveniste and Scheinkman theorem (see, eg, Theorems 4.10 and 9.10 in Stokey, Lucas and Prescott (1989)). □

Lemma S.3.12. 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 recursively 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 $(\boldsymbol{u}^{(t)}, \mathbf{w}^{(t)})_{t=0}^{\infty}$ and $(\boldsymbol{u}'^{(t)}, \mathbf{w}'^{(t)})_{t=0}^{\infty}$ denote the stochastic processes of menus generated by ξ^* starting from $\mathbf{v}$ and $\mathbf{v}'$, respectively. For each $t \in \mathbb{N}$, et

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

Thus, the process $\left(\boldsymbol{u}''^{(t)}, \mathbf{w}''^{(t)}\right)_{t=0}^{\infty}$ denotes the *history-wise* convex combination of the processes $(\boldsymbol{u}^{(t)}, \mathbf{w}^{(t)})_{t=0}^{\infty}$ and $(\boldsymbol{u}'^{(t)}, \mathbf{w}'^{(t)})_{t=0}^{\infty}$. Because ξ^* is feasible and delivers promises (ie, satisfies [**DP**], by virtue of Lemma 3.2) and because the constraint correspondence Γ has convex graph (by part (b) of Theorem 1), it follows that the process $(\boldsymbol{u}''^{(t)}, \mathbf{w}''^{(t)})_{t=0}^{\infty}$ 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 $T \in \mathbb{N}$ and two vectors $\mathbf{a}, \mathbf{b} \in \mathcal{U}^d$ with $\mathbf{a} \neq \mathbf{b}$ such that $\mathbf{P}(\boldsymbol{u}^{(T)} = \mathbf{a}) > 0$ and $\mathbf{P}(\boldsymbol{u}^{(T)} = \mathbf{b}) > 0$. Now, expanding the right-hand side of [**FE**] for T periods, we have

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

The strict inequality follows from convexity of $P(\cdot, s)$ (established in Lemma S.3.7), strict convexity of the $C(\cdot, i)$, and the fact that $\mathbf{P}(\boldsymbol{u}^{(T)} = \mathbf{a}) > 0$ and $\mathbf{P}(\boldsymbol{u}^{(T)} = \mathbf{b}) > 0$ where $\mathbf{a} \neq \mathbf{b}$. The weak inequality follows from the definition of $P(\cdot, s)$. This establishes strict convexity. □

Lemma S.3.13. Suppose that Condition R.1 holds. 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 $j = 2, \ldots, d$.

Proof. Consider part (a). Let $\mathbf{v} \in D$ and $s \in S$ be given, and define $\mathbf{v}' := (v_1 - \varepsilon, v_2, \ldots, v_d)$ 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 S.3.9). 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 1, Proposition A.11, and Lemma S.3.8 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 A.11). It is easy to see from the structure of that set that, for any $j \geq 2$ and fixed $\mathbf{v}_{-j} := (v_1, \ldots, v_{j-1}, v_{j+1}, \ldots, v_d)$, the set $\{y \in \mathcal{U} : (y, \mathbf{v}_{-j}) \in D\}$ is bounded. Non-monotonicity in the jth component then follows from Lemma S.3.8. □

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

Proof of Theorem 2. Lemmas S.3.5 and S.3.9 together imply that P satisfies [**FE**]. Convexity and continuous differentiability follow from Lemmas S.3.7 and S.3.11, respectively. Part (a) of the theorem follows from Lemma S.3.6. Part (b) of the theorem follows from Lemmas S.3.8 and S.3.13. For part (c), existence of a recursively optimal contract follows from Lemma S.3.10. The independence properties follow from part (a) of Lemma B.9 in Appendix B.3.3. Part (d) follows from Lemma S.3.9.

Finally, consider part (e). Point (i), strict convexity, follows from Lemma S.3.12. Given strict convexity and Lemma S.3.9, the policy correspondence Γ^* is singleton-valued, so there exists a unique conserving contract and hence at most one recursively optimal contract. By Lemma S.3.10, it is indeed the unique recursively optimal contract, establishing point (ii). Point (iii), continuity, follows from uniqueness and the upper hemi-continuity in Lemma S.3.9. Point (iv) follows from Lemma 3.2. □

S.4. A Useful Suboptimal Contract

In this appendix, we consider the case in which utility is of the CARA class and the Markov process satisfies either MLRP or PPR. Recall, from part (e) of Theorem 1, that in this case the largest recursive domain $D = V_d$. The goal is construct a (suboptimal) recursive contract $\zeta = (\zeta^f, \zeta^c)$ that is [**TVC**]-implementable starting from any $\mathbf{v} \in D$. This is then used to conclude that $D^* = V_d$ in part (f) of Theorem 1 (see Appendix A.3). We will also show that ζ generates well-defined and finite value for the principal starting from any $\mathbf{v} \in D$, a fact that is used to complete the proof of part (b) of Lemma 4.2 (see Appendix S.3.1). We note that this last piece is non-trivial because the flow cost function $C(\cdot)$ is unbounded both above and below.

We focus throughout on the $d = 2$ case; the construction in the general case is completely analogous 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 and the chain is fully connected by Assumption Markov, we have $\beta \in [0, 1)$. Because we are considering CARA utility $U(c) := -e^{-\rho c}$, 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.

S.4.1. Suboptimal Contracts

S.4.2. 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 S.4.1. 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. □

Lemma S.4.2. 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 S.4.1.

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. □

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

Lemma S.4.3. 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 S.4.1. But this is true because $\lambda^* \in (0, 1)$ and $\psi > 0$ on $(0, 1)$. □

Consider, again, the equations

$$\textbf{[PK}_1\textbf{]} \qquad \theta_1 x_1 + \alpha f_{11} w_{11} + \alpha f_{12} w_{12} = v_1$$

$$\textbf{[PK}_2\textbf{]} \qquad \theta_2 x_2 + \alpha f_{21} w_{21} + \alpha f_{22} w_{22} = v_2$$

$$\textbf{[IC}_{21}\textbf{]} \qquad \theta_2 x_1 + \alpha f_{21} w_{11} + \alpha f_{22} w_{12} \leq v_2$$

Multiply [**PK**$_1$] by $t_0 = \theta_2/\theta_1 < 1$ and subtract it from [**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

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

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

S.4.3. 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 S.4.4. 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 [**PK**$_2$], [**PK**$_1$], and [**IC**$_{21}$].

Proof. There exists $x_2 < 0$ such that $(x_2, \mathbf{v})$ satisfies [**PK**$_2$] 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 [**PK**$_1$] 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 [**IC**$^{\star}_{21}$] does not depend on the choice of (x_1, x_2). Thus, it suffices to show that [**IC**$^{\star}_{21}$] 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. □

S.4.4. Suboptimal Strategy — Part 2

Recall μ^* defined above in Appendix S.4.2, 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 S.4.5. 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 S.4.4, 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 S.4.3.) We can readily choose $\mathbf{w}_2 \in V^{(1)}$ and $x_2 < 0$ such that $\theta_2 x_2 + \mathbf{E}^{\mathbf{f}_2}[\mathbf{w}_2] = v_2$, ie, 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 [$\mathbf{IC}_{21}^{\star}$]. 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 [$\mathbf{IC}_{21}^{\star}$] (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. □

Let $(x_i(\mathbf{v}), \mathbf{w}_i(\mathbf{v}))_{i=1,2}$ be the solution described in Lemma S.4.5, 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.

S.4.5. 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 $s^{(1)}, \ldots, s^{(N)} \in S^N$, the continuation utility reaches W in at most N steps.

Lemma S.4.6. 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 [**PK$_2$**] 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 [**PK$_2$**] 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 [**PK$_1$**] and [**IC$^\star_{21}$**] can (respectively) be re-written as

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

(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,

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

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,

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

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 S.4.5 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. □

S.4.6. 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 S.4.7. 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 $Q^\zeta : V_2 \times S \to \mathbb{R}$.

Proof. The proof relies on Lemmas S.4.4, S.4.5, and S.4.6. Starting at $(\mathbf{v}, s)$, the contract defined in Lemmas S.4.5 and S.4.6 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 S.4.4. 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. □

S.5. Pathwise Properties of Markov Chains

In this section, we collect some miscellaneous facts about paths of Markov chains. Let $\mathcal{X}$ be the (countable) state space for a Markov process with transition probabilities $P(x, B)$ denoting the probability of transitioning from x to B. Let $\mathbf{P}$ denote the induced probability measure on the path space $\mathcal{X}^\infty$.

Lemma S.5.1. Let (X_n) be an $\mathcal{X}$-valued Markov process with transitions given by the kernel P, and suppose $x \in \mathcal{X}$ is *recurrent*. Then,

$$\mathbf{P}(X_n = x \text{ for infinitely many } n \mid X_0 = x) = 1$$

An elementary proof can be found on p. 577 of Shiryaev (1995). Now, to apply this result to our setting, let $\mathcal{X} := S$ and let $\mathbf{P}$ denote the measure on $\mathscr{H} = S^\infty$ induced by the type process defined in Section 2 (namely, in Assumption Markov).

Proposition S.5.2. The event $\{h \in \mathscr{H} : (s^{t-1}, s^t) = (i, j) \text{ for infinitely many } t\}$ occurs $\mathbf{P}$-a.s. for all $i, j \in S$.

Proof of Proposition S.5.2. Recall from Section 2 that $S = \{1, dots, d\}$ with transition probabilities $P(i, j) := f_{ij} > 0$.

It is useful to consider the *bivariate* Markov chain with states $B_S := \{(i, j) : i, j \in S\}$ and transition probabilities Q given by

$$Q\big((i, j), (k, \ell)\big) = \mathbf{1}\{j = k\}P(k, \ell)$$

where the indicator $\mathbf{1}\{j = k\} = 1$ if $j = k$ and 0 otherwise. Then, the two-step transition probabilities are given by

$$Q^{(2)}\big((i, j), (k, \ell)\big) = P(j, k)P(k, \ell) > 0$$

Therefore, all states communicate with each other, which implies that the Markov chain is indecomposable. But because the state space B_S is finite, by Theorem 1 (and the subsequent discussion) on p. 580 of Shiryaev (1995), at least one of the states must be recurrent. The indecomposability of the process then implies that all states are recurrent. An application of Lemma S.5.1 to the bivariate chain completes the proof. □

Corollary S.5.3. Let $\mathcal{J}_j := \{h \in \mathscr{H} : (s^{t-1}, s^t) = (d, j) \text{ infinitely often}\}$, and $\mathcal{J} := \bigcap_{j=1}^d \mathcal{J}_j$. Then, $\mathbf{P}(\mathcal{J}_j) = 1$ and hence $\mathbf{P}(\mathcal{J}) = 1$.

Proof. It follows from Proposition S.5.2 that $\mathbf{P}(\mathcal{J}_j) = 1$, and hence $\mathbf{P}(\mathcal{J}) = 1$. □

References

Atkeson, Andrew, and Robert E. Lucas Jr. 1992. 'On Efficient Distribution with Private Information'. *Review of Economic Studies* 59 (3): 427–453. (Cit. on pp. 4, 6).

Bloedel, Alexander W., R. Vijay Krishna and Oksana Leukhina. 2018. *Insurance and Inequality with Persistent Private Information*. Tech. rep. Stanford University. (Cit. on p. 1).

Farhi, Emmanuel, and Iván Werning. 2007. 'Inequality and Social Discounting'. *Journal of Political Economy* 115 (3): 365–402. (Cit. on pp. 5, 6).

Fernandes, Ana, and Christopher Phelan. 2000. 'A Recursive Formulation for Repeated Agency with History Dependence'. *Journal of Economic Theory* 91:223–247. (Cit. on p. 5).

Golosov, Mikhail, Aleh Tsyvinski and Nicolas Werquin. 2016. 'Recursive Contracts and Endogenously Incomplete Markets'. In *Handbook of Macroeconomics, Volume 2*. (Cit. on p. 3).

Green, Edward J. 1987. 'Lending and the Smoothing of Uninsurable Income'. In *Contractual Agreements for Intertemporal Trade*, ed. by Edward C. Prescott and Neil Wallace. University of Minnesota Press. (Cit. on pp. 4, 6).

Kamihigashi, Takashi. 2014. 'Elementary results on solutions to the Bellman equation of dynamic programming: existence, uniqueness, and convergence'. *Economic Theory* 56 (2): 251–273. (Cit. on p. 12).

Kreps, David M. 1977. 'Decision Problems with Expected Utility Criteria, I: Upper and Lower Convergent Utility'. *Mathematics of Operations Research* 2 (1): 45–53. (Cit. on p. 13).

Pavan, Alessandro, Ilya Segal and Juuso Toikka. 2014. 'Dynamic Mechanism Design: A Myersonian Approach'. *Econometrica* 82 (2): 601–653. (Cit. on p. 2).

Shiryaev, Albert N. 1995. *Probability*. Springer. (Cit. on p. 24).

Stokey, Nancy L., Robert E. Lucas Jr and Edward C. Prescott. 1989. *Recursive Methods in Economic Dynamics*. Cambridge, Ma: Harvard University Press. (Cit. on pp. 6, 12, 15, 16).

Thomas, Jonathan P., and Tim Worrall. 1990. 'Income Fluctuation and Asymmetric Information: An Example of a Repeated Principal-Agent Problem'. *Journal of Economic Theory* 51:367–390. (Cit. on pp. 5, 6).