CDE
August 2006

ON THE OPTIMAL COORDINATION OF UNINFORMED AGENTS BY AN INFORMED PRINCIPAL

Sudhir A. Shah
Email: sudhir@econdse.org

Delhi School of Economics
University of Delhi

Working Paper No. 147

Centre for Development Economics
Department of Economics, Delhi School of Economics

On the optimal coordination of uninformed agents by an informed principal

Sudhir A. Shah*

Department of Economics

Delhi School of Economics

March 30, 2006

We consider organizations with a single principal and many agents who interact in an environment with the following features: (a) Nature imperfectly informs the principal *via* a state-contingent signal, but not the agents, about the state of the world, (b) the principal selectively shares this information with the agents, thereby endogenously endowing them with private information that is coarser than his own, (c) the principal assigns action spaces to the agents, and (d) an agent's control over the choice from his assigned action space is inalienable. Designing an organization involves specifying (c) and specifying an information dissemination system for implementing (b). Searching for an optimal design involves (1) deriving optimal performance from each design, and (2) comparing designs on the basis of their best performances. Our existence results show the feasibility of performing Step (1) in a large class of cases.

JEL classification: C62, D02, D23, D82, L23

Key words: Existence theorems, optimal design, team, organization, principal-agent model

* Address: Department of Economics, Delhi School of Economics, University of Delhi, Delhi 110007, India. Tel: (+91)(11) 27667005. Fax: (+91)(11) 27667159. Email: sudhir@econdse.org

1. Introduction

1.1 Motivation

The formal study of organizations involves two significant classes of complications: (a) structural complexity, and (b) complexities stemming from asymmetries of information and incentives. (a) refers to the myriad ways of assigning rights and responsibilities, and structuring relationships of power and authority, among the members of the organization, e.g., a given member might serve as a common principal (resp. agent) for various agents (resp. principals) or may perform multiple roles by acting as a principal in one relationship and an agent in another. (b) refers to the problems of inducing members to make decisions in the organization's interest when they possess private information or their decisions cannot be monitored effectively and contracted upon. Depending on the context, these complications may be treated as descriptors of the environment, i.e., as primitives of a model, or as aspects of the organization design chosen in response to the environment, i.e., as constructions chosen by a rational decision-maker. An objective of the economic theory of organizations[1] is to solve the design problem: given an environment, a class of feasible designs, and an evaluation criterion, find an optimal organization design.[2]

In order to solve the design problem for a given environment, the first step is to use the given criterion to evaluate the feasible organization designs in terms of the outcomes that follow from them. In this paper, we model this problem and show that it can be solved in the following framework. The environment that we consider has the following salient features: (a) Nature initially endows only one member of the organization with private information, and (b) the informed member is the sole principal and the other members are his agents. Feature (b) will be formally represented by identifying the organization's evaluation criterion with the principal's payoff function. The class of organization designs considered in response to this environment is restricted by the following features: (a) the division of labor in the organization is such that the principal designs the organization and coordinates activities, while the agents are delegated the various tasks that constitute the organization's activities and determine the payoff-relevant outcomes, and (b) the principal uses two channels to communicate with the agents, one private and the other public. With respect to (a), an agent's performance of assigned tasks is to be guided solely by his incentives. With respect to (b), the private channel is used to share, privately and selectively, portions of the principal's information with the agents, thereby endowing them with private information that is coarser than the principal's. The public channel is used to publicly recommend action choices to the agents.

Thus, we consider organizations that operate sequentially as follows: (1) Nature informs the principal, but not the agents, about the state of the world, (2) the principal selectively and privately shares this information with the agents, and (3) an agent chooses an action after hearing the publicly issued action recommendation by the principal. Every member's payoff will depend on the state and the profile of actions chosen by the agents. When deciding whether to obey the principal, an agent's belief regarding the state will be conditioned on the private signal received in Step (2) and the public recommendation received in Step (3). Our notion of equilibrium will require the principal to issue recommendations that every agent will choose to obey, assuming that the other agents will do so too.

The principal's role in the postulated organization is twofold: (α) to design the organization, and (β) to use this design and his information to guide the agents' decision-making. Role (α) is performed before seeing Nature's state-contingent signal, i.e., prior to Step (1) in the scheme described above. It involves specifying the agents' action spaces and the organization's communication system. Role (β) involves recommending actions to the agents in the context of the chosen organization design. The action spaces serve as the language in which the principal issues recommendations to the agents and as the sets from which the agents choose their responses in Step (3). The organization's communication system determines the extent of information sharing by the principal in Step (2). The action recommendations in Step (3) are an indirect source of information for the agents as they may reveal information about the state in addition to what is shared directly with them in Step (2).

An agent's role in the organization is to choose an action from the action space assigned to him by the design after receiving his share of the principal's information and the principal's action recommendation. The principal uses his privileged position in the organization to manipulate the agents' incentives, *via* the design and the recommended actions, in order to induce action choices that are optimal from his point of view subject to the constraint that the action chosen by each agent must be incentive compatible for him given the constraints imposed by the organization design.

The objective of this paper is to model Role (β) and show that the principal can perform this role optimally. More precisely, given an organization design, we show the existence of a mapping that generates state-contingent action recommendations to the agents with the following properties. First, it respects the principal's information constraint *via* an appropriate measurability requirement. Secondly, it induces obedience by all the agents in every state. Thirdly, among all mappings possessing the above two properties, the given mapping is optimal from the principal's point of view. This exercise enables the principal to evaluate a particular organization design prior to receiving his private information, thus defining a value function over the set of organization designs. We show that this *ex ante* evaluation program can be carried out in two distinct environments; see the hypotheses of Theorems 2.4.4 and 2.4.5.

Given the value function generated by a solution of Role (β), the choice of a design in Role (α) is now a matter of maximizing the derived value function over the set of feasible designs. In order to get quickly to the formal results, we relegate our observations regarding Role (α) to Section 3. However, a reader wishing to contextualize the above discussion of organizations may make a detour to Section 3.2 where we describe the problem of "designing a firm" as an example of the general problem of designing an organization.

1.2 The literature

The organization described above resonates with the notions of teams described in [20] and syndicates described in [31]; in particular, it has a number of congruencies with the generalized team described in Part Two of [20]. For instance, unlike a classical team, but like a generalized team, our organization will allow for heterogeneous preferences. Apart from these seminal sources, our work is related to the vast literature on principal-agent problems with a single principal.

The literature on organizations has considered various selections from the above-mentioned complexities to describe the "environment" and a variety of classes of organization "designs". With respect to structural complexity, while multiple role-playing is not modelled generally and models with a single agent facing multiple principals are of relatively recent vintage (e.g., [5] and [9]), the bulk of the literature concerns simple organizations consisting of a single principal facing one or more agents. With respect to complications arising from informational asymmetries, while some models endow the principal with private information (e.g., [21], [22] and [25]), most models allow only the agents to have private information.

We start with the principal-agent literature that does not allow the principal to have private information. The models in this literature feature a selection from the following sequence of events: (a) Nature privately conveys to each agent his type, (b) agents privately report their types, possibly falsely, to the principal, (c) the principal sends privacy-preserving action recommendations to the agents based on the received type reports, and (d) each agent chooses an action based on his type and the action recommendation received from the principal. Note that (b) and (c) feature communication that is modelled using the "direct" message spaces. The principal's problem is to choose privacy-preserving action recommendations, conditional on the type-reports, that maximize his expected payoff. This problem is simplified further by restricting the principal's attention to reports-to-recommendations mappings that, in a non-cooperative fashion, induce truthful reporting of types and obedient choice of actions by the agents.[3]

This class of models can be further divided into three sub-classes. First, there are the pure "hidden characteristic" models, i.e., models that suppress steps (c) and (d) of the general model described above and replace them with an outcome function that directly maps agents' type reports to outcomes; we refer the reader to Chapter 7 of [13] for an introduction to this voluminous literature. In effect, each agent has finer exogenously specified information than the principal and an agent's control over his action choice (if any) is alienated. Secondly, there are the pure "hidden action" models, i.e., models that suppress steps (a) and (b) of the general model described above; examples include [14], [16], [19], [23] and [29]. In effect, neither the principal nor the agents have exogenously specified private information. Thirdly, there are models that allow agents to have exogenously specified private information and retain control over their actions; for instance, see [18].

The model we study differs from this entire literature in two essential and obvious ways. First, in our model, the principal has private information that he distributes among the agents. Therefore, each agent's information regarding the state is coarser than that of the principal, rather than the other way around. Secondly, in our model, the principal communicates publicly with the agents rather than privately. Moreover, our model departs from the pure hidden characteristic models in that it does not alienate an agent's control over his action choice. It also departs from the pure hidden action models in that it features informational asymmetries in addition to those related to the observability of endogenously generated action choices by the agents.

We now turn to the second strand of the literature on principal-agent models, i.e., models in which the principal has private information, as in our model. This literature includes [25] in a largely cooperative game setting and [21] and [22] in the non-cooperative game setting. The motivation for this literature can be illustrated by the dilemma faced by the seller of a used car (the principal) after learning the quality of the car (the principal's private information): if it is common knowledge that he may offer a warranty to the buyer, then his (non-)offer of a warranty itself conveys information to a potential buyer regarding the quality of the car. More generally, the contract offered by an informed principal may leak his private information. The papers cited above study the resulting problem: what is the optimal way for the principal to structure the contract when the contract may reveal some of his private information? As the common values version of the problem studied in [22] has more in common with our setting, all further discussion will cite only this paper as Maskin-Tirole.

The Maskin-Tirole model is as follows. Consider a principal contracting with a single agent. The principal has private information while the agent does not. The model has three stages after the principal learns his type. First, the principal offers a contract to the agent. Next, the agent either rejects that contract, thereby ending the game, or accepts it, thereby sending the play of the game to the third stage; this is the only decision made by the agent in this model. If the offered contract is accepted in the second stage, then it is implemented in the third stage. A contract maps the principal's type to the outcome space.[4] The principal's ability to offer *any* such mapping as a contract reflects the fact that the agent does not control any outcome-relevant action choice when the contract is implemented; otherwise, the class of contracts from which the principal makes a choice would be limited to those that induce appropriate action choices by the agent.

The model studied in this paper may be seen as an enrichment of the third stage of the Maskin-Tirole model in the following dimensions. First, we introduce action choice by the agent as an ingredient of the implementation stage. The implied moral hazard problem restricts the set of implementable contracts that the principal can offer in the first stage of the Maskin-Tirole model. This restriction is in the nature of a refinement and the implications of this refinement for the full three-stage Maskin-Tirole results are unclear. Secondly, we allow multiple agents in our model. This is a non-trivial complication as these agents have heterogeneous preferences and information, and control action choices. Thirdly, we allow direct information revelation by the principal instead of the indirect way postulated in the Maskin-Tirole model. As the principal controls the process of information revelation, the agents are endowed with private information, albeit coarser than the principal's information.

Returning to the example of the car salesman, the seller may reveal information to potential buyers not just indirectly *via* the (non-)offer of a warranty, but also directly, say *via* the selective sharing of maintenance records or technical information certified by an "honest broker". More generally, we have the question that characterizes the principal's Role (α): how should the principal use his ability to structure the information dissemination and action allocation in an organization? Obviously, the same question arises in signalling models (e.g., [7]), especially if there are multiple signal receivers.

1.3 The model

We begin the modelling of the principal's problem with respect to Role (β) by describing the data that is exogenously given. $(\Omega, \mathcal{F})$ is a measurable state space. $(T, \mathcal{T})$ is the measurable space of signals received by the principal from Nature. $P \in \Delta(\Omega)$ is the principal's prior belief about the state; henceforth, $\Delta(Y)$ will denote the set of probability measures on a measurable set Y. The principal's information about the state is derived from the signal generated by the surjective measurable function $s : \Omega \to T$. $\Lambda : T \times \mathcal{F} \to [0, 1]$ is the transition kernel generating the principal's belief about the state conditional on the signal received from Nature. Interpreting Λ as a regular conditional distribution on $(\Omega, \mathcal{F})$ imposes the following restrictions on it: (a) for every $t \in T$, $\Lambda(t, .) \in \Delta(\Omega)$, (b) for every $E \in \mathcal{F}$, $\Lambda(., E)$ is measurable, (c) $P(.) = \int_T P \circ s^{-1}(dt)\Lambda(t, .)$, and (d) $\Lambda(t, s^{-1}(\{t\})) = 1$ for $P \circ s^{-1}$ almost every $t \in T$.[5] $(O, \mathcal{O})$ is the measurable outcome space. $w : O \to \Re$ is the principal's von Neumann-Morgenstern utility, representing his preference $\succeq$ defined on $\Delta(O)$. I is the set of agents. $v_i : O \to \Re$ is agent i's von Neumann-Morgenstern utility function, representing agent i's preference $\succeq_i$ on $\Delta(O)$.

Now consider an organization design $d = \{((T_i, \mathcal{T}_i), s_i, (A_i, \mathcal{A}_i)), \sigma \mid i \in I\}$; the problem of choosing an optimal design d defines the principal's Role (α), which we consider in Section 3. $(T_i, \mathcal{T}_i)$ is the measurable space of signals received by agent i from the principal. $s_i : T \to T_i$ is the surjective measurable mapping that generates agent i's signal as a function of the principal's information. The collection $\{(T_i, \mathcal{T}_i), s_i \mid i \in I\}$ is the principal's information dissemination system. $(A_i, \mathcal{A}_i)$ is agent i's measurable action space. $A = \prod_{i \in I} A_i$ is the set of action profiles and $\sigma : \Omega \times A \to O$ is the measurable outcome function. Combining design d with the exogenously given data, we have the data

$$\Gamma = \{(\Omega, \mathcal{F}), (T, \mathcal{T}), (P, s, \Lambda), (O, \mathcal{O}), w, I, (v_i, (T_i, \mathcal{T}_i), s_i, (A_i, \mathcal{A}_i))_{i \in I}, \sigma, \lambda\} \tag{1.3.1}$$

where $\lambda \in \Delta(\Omega \times A)$ is the agents' common prior belief about the state and the action profile to be recommended by the principal.

Agent i's information is generated by the mapping $F_i : \Omega \times A \to T_i \times A$, where $F_i(\omega, a) = (s_i \circ s(\omega), a)$. F_i has distribution $\lambda \circ F_i^{-1}$ on $T_i \times A$. Let $\Lambda_i : T_i \times A \times \mathcal{F} \to [0, 1]$ generate agent i's posterior belief about the state conditional on receiving the message generated by F_i. This requires that (a) $\Lambda_i(t_i, a, .) \in \Delta(\Omega)$ for every $(t_i, a) \in T_i \times A$, (b) $\Lambda_i(., ., E)$ is measurable for every $E \in \mathcal{F}$, and (c) $\lambda(. \times A) = \int_{T_i \times A} \lambda \circ F_i^{-1}(dt_i \times da)\Lambda_i(t_i, a, .)$. Of the message $F_i(\omega, a)$ received by agent i, the first component $s_i \circ s(\omega)$ is private, while the second component a is public, with the interpretation that a_i is the action choice being recommended to agent i. Define $u_i : T \times A^2 \to \Re$ by

$$u_i(t, a, b) = \int_\Omega \Lambda_i(s_i(t), a, d\omega) v_i \circ \sigma(\omega, b) \tag{1.3.2}$$

$u_i(t, a, b)$ is agent i's expected utility if the principal receives the signal $t \in T$ (whereupon agent i receives the private signal $s_i(t)$), publicly recommends the action profile $a \in A$ and the chosen action profile is $b \in A$. Note that agent i's expected utility does not depend directly on t, but only indirectly *via* agent i's signal $s_i(t)$.

Define $B_i : T \times A \rightrightarrows A_i$ by $B_i(t, a) = \cap_{c \in A_i}\{b \in A_i \mid u_i(t, a, a_{-i}, b) \geq u_i(t, a, a_{-i}, c)\}$, which is the set of optimal action choices by agent i if the principal receives signal t, recommends the action profile $a \in A$, and the other agents choose actions as *per* the principal's recommendation. Define $B : T \times A \rightrightarrows A$ by $B(t, a) = \prod_{i \in I} B_i(t, a)$. Also define $W : T \times A \to \Re$ by

$$W(t, a) = \int_\Omega \Lambda(t, d\omega) w \circ \sigma(\omega, a) \tag{1.3.3}$$

which is the principal's expected utility conditional on receiving signal $t \in T$ from Nature and action profile $a \in A$ being implemented by the agents.

Definition 1.3.4. *(A) $a : T \to A$ is an equilibrium coordination plan (ECP) for Γ if*

(a) a is measurable, and

(b) $a(t) \in B(t, a(t))$ for every $t \in T$.

Let $\mathcal{E}$ be the set of ECPs for Γ.

(B) An ex ante optimal ECP for Γ is $a \in \mathcal{E}$ such that $\int_\Omega P(d\omega) w \circ \sigma(\omega, a \circ s(\omega)) \geq \int_\Omega P(d\omega) w \circ \sigma(\omega, b \circ s(\omega))$ for every $b \in \mathcal{E}$.

(C) An ex post optimal ECP for Γ is $a \in \mathcal{E}$ such that $W(t, a(t)) \geq W(t, b(t))$ for every $t \in T$ and every $b \in \mathcal{E}$.

In the definition of an ECP, (a) ensures that the coordination plan respects the principal's information constraint, while (b) ensures that, for every signal t received by the principal, his action recommendations $a(t)$ induce obedience by all the agents. An *ex ante* optimal ECP yields the principal the highest expected payoff among all ECPs before receiving Nature's signal. However, there is no guarantee that the principal would choose to implement an *ex ante* optimal ECP after receiving Nature's signal; an *ex post* optimal ECP guarantees this. The following result provides the natural link between *ex ante* and *ex post* optimal ECPs.

Lemma 1.3.5. *Consider Γ. If $w \circ \sigma$ is integrable and $a \in \mathcal{E}$ is an ex post optimal ECP, then a is an ex ante optimal ECP.*

Given this fact, we will search for *ex post* optimal ECPs and henceforth drop the qualifier *"ex post"*. Our method of showing the existence of an optimal ECP is as follows. First, we characterize an ECP as a measurable selection from an appropriate equilibrium mapping from T to A (Lemmas 2.3.1, 2.3.2 and 2.3.4). Next, we compose the equilibrium mapping with W to derive a mapping from T to $\Re$; for each $t \in T$, this mapping yields the set of payoffs for the principal resulting from the set of ECPs. Finally, an implicit measurable selection theorem is used to extract an ECP from the equilibrium mapping that, for every $t \in T$, yields the highest possible payoff to the principal among all the ECPs.

Some interpretative remarks regarding Γ are appropriate at this stage.

First, assuming a common belief λ across agents is a matter of convenience and is not required for our results. Moreover, although we have taken λ as primitive data in the interest of descriptive economy, it may itself be derived from other specifications of primitive data. For instance, suppose $\mathcal{M}$ is a set of mappings $a : T \to A$ that may generate the principal's action recommendations contingent on his information. Let $Q \in \Delta(\Omega \times \mathcal{M})$ be the agents' belief regarding the state and the mapping to be used by the principal to generate recommendations. Define $\xi : \Omega \times \mathcal{M} \to \Omega \times A$ by $\xi(\omega, a) = (\omega, a \circ s(\omega))$. Given this formalism, we may set $\lambda = Q \circ \xi^{-1}$. In particular, if the agents share the principal's belief P regarding the state and believe that the mapping $a \in \mathcal{M}$ is used to generate the action recommendations, then $Q = P \times \delta_a$ and $\lambda = (P \times \delta_a) \circ \xi^{-1}$.

Secondly, while the principal needs to know all the data in Γ in order to formulate his decision problem and find an optimal ECP, the agents do not necessarily need to know all of Γ in order to formulate their decision problems as *per* our equilibrium concept. For instance, the agents do not need to know w and (P, Λ). Moreover, agent i need not know $(v_j, (T_j, \mathcal{T}_j), s_j)$ for $j \in I - \{i\}$.

Thirdly, while the definition of an ECP does not explicitly feature participation constraints for the agents, the provision of appropriate action spaces and outcome functions implicitly allows such constraints to be incorporated. For instance, an agent's "participation decision" can be captured implicitly in our framework by allowing him an action corresponding to "zero effort" and rigging the outcome function so that the resulting outcome is no worse than an agent's "reservation outcome". Naturally, such requirements will place restrictions on the class of feasible organization designs.

1.4 An application

Consider a firm with a single owner and a set of workers I. Let $O = O_0 \times \prod_{i \in I} O_i$ be the outcome space, with O_i as worker i's private outcome space (e.g., the space of i's wages) and O_0 as the supplementary outcome space (e.g., the space of the owner's revenues). Worker i's utility is generated by $v_i^* : O_i \to \Re$, i.e., $v_i = v_i^* \circ \pi_i$, where π_i projects O on O_i. The other aspects of the environment are as described in Section 1.3. We may interpret a state $\omega \in \Omega$ as representing technology and $s(\omega) \in T$ as the owner's information regarding the state.

The firm's design is described as follows. The action spaces $\{A_i \mid i \in I\}$ and the mappings s_i are as described in Section 1.3. We set $T_i \subset O_i^A$, i.e., the information shared by worker i with the owner is a mapping $t_i : A \to O_i$, which we interpret to be a contract that generates worker i's private outcome as a function of the action profile; worker i's wage may depend not only on his own effort but also on i's effort relative to the efforts of other workers, as in a tournament set-up studied in [19]. So, if the state is ω, then the owner will receive signal $s(\omega)$ and award contract $s_i \circ s(\omega)$ to worker i. Finally, define the outcome function $\sigma : \Omega \times A \to O$ by

$$\sigma(\omega, a) = \sigma_0(\omega, a) \times (s_i \circ s(\omega)(a))_{i \in I}$$

Using the above definitions of v_i^* and σ and the regularity of Λ_i, if t is the signal received by the owner, a is the profile of actions recommended by the owner and b is the profile of implemented actions, then worker i's payoff specializes from (1.3.2) to

$$\begin{aligned} u_i(t, a, b) &= \int_\Omega \Lambda_i(s_i(t), a, d\omega) v_i^* \circ \pi_i \circ \sigma(\omega, b) \\ &= \int_\Omega \Lambda_i(s_i(t), a, d\omega) v_i^* \circ s_i \circ s(\omega)(b) \\ &= v_i^* \circ s_i(t)(b) \end{aligned} \tag{1.4.1}$$

The nature of preferences and the outcome function imply that the only aspect of state ω that is relevant to worker i is the contract $s_i \circ s(\omega)$ awarded to him. Consequently, the regularity of Λ_i implies that $s_i \circ s(\omega) = s_i(t)$ almost surely, which enables the simplification evident in (1.4.1). In the general model, the privateness of each agent's signal is a means for manipulating agents' incentives *via* their posterior beliefs regarding the state. As is clear from (1.4.1), the privateness of signal $s_i(t)$ is irrelevant in this application. Given this set-up and design, does the owner have an optimal ECP? We show in Section 2.5 that the nature of $s_i(t)$ is crucial for guaranteeing an affirmative answer as it directly and exclusively determines the dependence of i's private outcome and payoff on the profile of actions. More precisely, we provide two classes of contracts that allow the existence of an optimal ECP.

To summarize, the application involves the specialization of some of the data describing Γ, namely, the outcome space O, the agents' payoff functions v_i, the agents' signal spaces T_i and the outcome function σ. This specialized version of Γ will be denoted by Γ^*. In Section 2.5, we provide assumptions regarding Γ^* that enable the application of the general results regarding Γ.

1.5 Plan of paper

The remainder of the paper is organized as follows. Section 2 is devoted to the statements and proofs of various results concerning the existence of optimal ECPs. In Section 2.1, we state various mathematical conventions that we employ in this paper. Section 2.2 is devoted to deriving the properties of the expected utility functions u_i and W, defined by (1.3.2) and (1.3.3) respectively, under a variety of assumptions regarding the primitive data Γ. Although these properties play a vital role in the existence results related to ECPs (resp. optimal ECPs) of Section 2.3 (resp. Section 2.4), they may be taken as given if the reader wishes to pass directly to Section 2.3. In Section 2.5, we apply the results of Sections 2.3 and 2.4 to the application described in Section 1.4. We consider the problem of choosing an optimal organization design in Section 3 and conclude in Section 4. The proofs are relegated to the two appendices.

2. Solutions of existence problem

2.1 Conventions

We shall use the following conventions without specific comment: (1) a measurable space refers to a pair $(X, \mathcal{X})$, where X is a set and $\mathcal{X}$ is a σ-algebra on X; (2) subsets of measurable (resp. topological) spaces are given the trace σ-algebra (resp. subspace topology); (3) products of measurable (resp. topological) spaces are given the product σ-algebra (resp. topology); (4) if X is a topological space, then it is given the Borel σ-algebra, denoted by $\mathcal{B}(X)$; (5) the set of real numbers, $\Re$, is given the Euclidean metric topology; (6) if X and Y are topological spaces, then the space of continuous functions $f : X \to Y$ is denoted by Y^X; (7) if X is topological, then the space of bounded elements of $\Re^X$ is denoted by $\mathcal{C}(X)$; (8) if $(X, \mathcal{X})$ and $(Y, \mathcal{Y})$ are measurable spaces, then the product measurable space is denoted by $(X \times Y, \mathcal{X} \times \mathcal{Y})$. Lemma B.1 implies that conventions (2) and (4) are consistent. Lemma B.2 implies that conventions (3) and (4) are consistent for a wide class of situations relevant to this paper. Given a measurable space $(X, \mathcal{X})$ and topological spaces Y and Z, a function $f : X \times Y \to Z$ is said to be Caratheodory if $f(., y)$ is measurable for every $y \in Y$ and $f(x, .)$ is continuous for every $x \in X$; by convention, whenever we describe a function as being "Caratheodory", we mean that it is measurable in the first argument and continuous in the second argument. We shall abbreviate "locally convex topological vector space" to l.c.s.

$F : X \rightrightarrows Y$ denotes a set-valued mapping with domain X and values in 2^Y. Given topological spaces X and Y, F is said to be upper (resp. lower) hemicontinuous if $F^+(E) = \{x \in X \mid F(x) \subset E\}$ (resp. $F^-(E) = \{x \in X \mid F(x) \cap E \neq \emptyset\}$) is open in X for every E open in Y. If $(X, \mathcal{X})$ is a measurable space and Y is a topological space, then F is said to be measurable (resp. weakly measurable) if $F^-(E) \in \mathcal{X}$ for every E closed (resp. open) in Y. We note some well-known facts that we shall use repeatedly without explicit reference.

Remark 2.1.1. *A compact metric space is second countable ([10], Theorem XI.4.1), separable ([10], Theorem VIII.7.3) and complete ([10], Corollary XIV.2.4); consequently, it is a Polish space, and therefore a Souslin space (see [15]). Given convention (2), a product of compact spaces is compact ([10], Theorem XI.1.4); a countable product of metrizable spaces is metrizable ([10], Corollary IX.7.3).*

2.2 Properties of the expected utility function

In this section we establish conditions on Γ that ensure u_i and W are either continuous or Caratheodory. We prove the results related to u_i but omit the proofs of results related to W as they are analogous.

The properties of u_i, defined by (1.3.2), are ingredients in the existence theorems proved in Sections 2.3 and 2.4. Lemma 2.2.1 specifies assumptions on Γ guaranteeing that

u_i is continuous. Lemmas 2.2.2 and 2.2.3, and Corollary 2.2.4, provide conditions guaranteeing that u_i is Caratheodory, i.e., $u_i(., a, b)$ is measurable and $u_i(t, ., .)$ is continuous.

The common features of all three lemmas are: A and O are topological spaces and v_i is continuous. The state space Ω is given topological structure in Lemmas 2.2.1 and 2.2.2, so that $\mathcal{F} = \mathcal{B}(\Omega)$, but not in Lemma 2.2.3. The outcome function σ is assumed to be continuous in Lemmas 2.2.1 and 2.2.2, but this requirement is weakened in Lemma 2.2.3. Finally, for $E \in \mathcal{F}$, while $\Lambda_i(., ., E)$ is assumed to be continuous in Lemma 2.2.1, it is required to be only a Caratheodory mapping in Lemmas 2.2.2 and 2.2.3.

Lemma 2.2.1. *Consider Γ and $i \in I$. If*

(a) Ω is compact metric; A, O, T and T_i are topological,

(b) σ, v_i and s_i are continuous, and

(c) $\Lambda_i : T_i \times A \times \mathcal{B}(\Omega) \to [0, 1]$ is such that $\Lambda_i(t_i, a, .) \in \Delta(\Omega)$ for every $(t_i, a) \in T_i \times A$ and $\Lambda_i(., ., E)$ is continuous for every $E \in \mathcal{B}(\Omega)$,

then u_i is continuous.

The next two results provide conditions that ensure u_i is a Caratheodory mapping, as required in Theorem 2.3.5.

Lemma 2.2.2. *Consider Γ and $i \in I$. If*

(a) Ω is compact metric; A and O are topological,

(b) σ and v_i are continuous, and

(c) $\Lambda_i : T_i \times A \times \mathcal{B}(\Omega) \to [0, 1]$ is such that $\Lambda_i(t_i, a, .) \in \Delta(\Omega)$ for every $(t_i, a) \in T_i \times A$ and $\Lambda_i(., ., E)$ is Caratheodory for every $E \in \mathcal{B}(\Omega)$,

then u_i is Caratheodory.

The next result weakens the continuity assumption on σ to a property intermediate between continuity and the Caratheodory property. Denote by $B(\Omega, \mathcal{F})$ the space of functions that are uniform limits of finite linear combinations of functions drawn from $\{1_E \mid E \in \mathcal{F}\}$; equip $B(\Omega, \mathcal{F})$ with the supremum norm.

Lemma 2.2.3. *Consider Γ and $i \in I$. If*

(a) A and O are topological,

(b) σ is Caratheodory and v_i is continuous,

(c) $a \mapsto v_i \circ \sigma(., a)$ is a continuous mapping from A to $B(\Omega, \mathcal{F})$, and

(d) $\Lambda_i : T_i \times A \times \mathcal{F} \to [0, 1]$ is such that $\Lambda_i(t_i, a, .) \in \Delta(\Omega)$ for every $(t_i, a) \in T_i \times A$

and $\Lambda_i(.,.,E)$ is Caratheodory for every $E \in \mathcal{F}$,
then u_i is Caratheodory.

Corollary 2.2.4. *Given (a) and (d), if Ω and A are compact, and (b) is replaced by*

(b′) σ and v_i are continuous,

then u_i is Caratheodory.

The properties of W, defined by (1.3.3), are ingredients in the existence theorems proved in Section 2.4. Lemma 2.2.5 specifies assumptions on Γ guaranteeing that W is continuous. It is proved by mimicking the proof of Lemma 2.2.1.

Lemma 2.2.5. *Consider Γ. If*

(a) Ω is compact metric; A, O and T are topological,

(b) σ and w are continuous, and

(c) $\Lambda : T \times \mathcal{B}(\Omega) \to [0,1]$ is such that $\Lambda(t,.) \in \Delta(\Omega)$ for every $t \in T$ and $\Lambda(.,E)$ is continuous for every $E \in \mathcal{B}(\Omega)$,

then W is continuous.

Lemmas 2.2.6 provides conditions guaranteeing that W is Caratheodory. It is proved by mimicking the proof of Lemma 2.2.3.

Lemma 2.2.6. *Consider Γ. If*

(a) A and O are topological,

(b) σ is Caratheodory and w is continuous, and

(c) $\Lambda : T \times \mathcal{F} \to [0,1]$ is such that $\Lambda(t,.) \in \Delta(\Omega)$ for every $t \in T$ and $\Lambda(.,E)$ is measurable for every $E \in \mathcal{F}$,

then W is Caratheodory.

2.3 Existence of equilibrium coordination plans

In this section we provide sufficient conditions for the existence of an ECP for Γ. The following lemma is the key to Theorems 2.3.3 and 2.4.4, which are existence results for ECPs and optimal ECPs respectively.

Lemma 2.3.1. *Given Γ, if*

(a) I is countable,

(b) T is a compact Hausdorff space, and

for every $i \in I$,

(c) A_i is a convex, compact and metrizable subset of an l.c.s. L_i, and

(d) u_i is continuous and $u_i(t, a, b_{-i}, .) : A_i \to \Re$ is quasi-concave for every $(t, a, b_{-i}) \in T \times A \times A_{-i}$,

then

(A) B is upper hemicontinuous, with nonempty, convex and compact values.

Define $\Xi : T \rightrightarrows A$ by $\Xi(t) = \{a \in A \mid a \in B(t, a)\}$. Then,

(B) Ξ has nonempty compact values and $\mathrm{Gr}\,\Xi$ is closed in $T \times A$,

(C) Ξ is (weakly) measurable, and

(D) $a : T \to A$ is an ECP for Γ iff. a is a measurable selection from Ξ.

The following lemma is a slight variation on Lemma 2.3.1.

Lemma 2.3.2. *Given Γ, if Assumptions (a), (c) and (d) of Lemma 2.3.1 are satisfied and*

(b′) T is a separable metric space and $(T, \mathcal{B}(T))$ is a complete measurable space, then the conclusions of Lemma 2.3.1 hold.

Using the above results, we have an existence theorem for ECPs that exploits the topological structure of T and the continuity of each expected utility function u_i. See Lemma 2.2.1 for conditions on Γ that ensure the continuity of u_i.

Theorem 2.3.3. *If Γ satisfies the assumptions of Lemma 2.3.1 (resp. 2.3.2), then there exists an ECP for Γ.*

The second existence result dispenses with the topological structure of T and the continuity of the expected utility functions, with the former replaced by measure-theoretic requirements and the latter replaced by the requirement of a Caratheodory expected utility function. We prepare for this result by proving the following lemma, which is the key to Theorems 2.3.5 and 2.4.5, which are existence results for ECPs and optimal ECPs respectively.

Lemma 2.3.4. *Given Γ, if*

(a) I is countable,

(b) $(T, \mathcal{T})$ is a complete measurable space, and

for every $i \in I$,

(c) A_i is a convex compact subset of a separable Banach space L_i with $\mathrm{Int}\, A_i \neq \emptyset$,[6] *and*

(d) $u_i : T \times A^2 \to \Re$ is Caratheodory and $u_i(t, a, b_{-i}, .) : A_i \to \Re$ is quasi-concave for every $(t, a, b_{-i}) \in T \times A \times A_{-i}$,

then there exists a mapping $\phi : T \times A \Rightarrow A$ such that

(A) ϕ has nonempty, compact and convex values,

(B) ϕ is weakly measurable,

(C) $\phi(t, .)$ is upper hemicontinuous for every $t \in T$, and

(D) $a : T \to A$ is an ECP for Γ iff. a is a measurable function such that $a(t) \in \phi(t, a(t))$ for every $t \in T$.

Define $\Phi : T \Rightarrow A$ by $\Phi(t) = \{a \in A \mid a \in \phi(t, a)\}$. Then,

(E) Φ has nonempty closed values,

(F) Φ is measurable, and

(G) $a : T \to A$ is an ECP for Γ iff. a is a measurable selection from Φ.

We immediately have the required result.

Theorem 2.3.5. *If Γ satisfies the assumptions of Lemma 2.3.4, then there exists an ECP for Γ.*

2.4 Existence of optimal equilibrium plans

We now complete our program by providing sufficient conditions for the existence of (*ex post*) optimal ECPs. Define $G : T \Rightarrow \Re$ by $G(t) = W(\{t\} \times \Xi(t))$ and $g : T \Rightarrow \Re$ by $g(t) = \sup G(t)$. Given $t \in T$, $G(t)$ is the set of conditional expected utilities for the principal that can result from some ECP and $g(t)$ is the supremum of these utilities. Define $C : T \Rightarrow A$ by $C(t) = W(t, .)^{-1}([g(t), \infty)) \cap \Xi(t)$. Given $t \in T$, $C(t)$ is the set of equilibrium action profiles that maximize the principal's expected utility.

Lemma 2.4.1. *Let Γ satisfy the assumptions of Lemma 2.3.1. If W is Caratheodory, then*

(A) g is a measurable selection from G, and

(B) there exists a measurable selection a from Ξ such that $g(t) = W(t, a(t))$ for every $t \in T$; equivalently, there exists a measurable selection a from C.

Lemma 2.4.2. *Let Γ satisfy the assumptions of Lemma 2.3.2. If W is Caratheodory, then conclusions (A) and (B) of Lemma 2.4.1 hold. Moreover, C is measurable with nonempty closed values.*

Continuity assumptions on W yield richer properties of g and C.

Lemma 2.4.3. *Suppose* Γ *satisfies the assumptions of Lemma 2.3.1 (resp. 2.3.2).*

(A) If W *is upper semicontinuous, then* g *is an upper semicontinuous selection from* G*, and consequently,* g *is measurable.*

(B) If T *is compact metric and* W *is continuous, then there exists* $T_0 \subset T$ *such that*

(i) $T_0 = \cap_{n \in \mathcal{N}} T_n$*, where each* T_n *is open and dense in* T*,*

(ii) T_0 *is a Baire space of second category that is dense in* T*, while* $T - T_0$ *is a set of first category;* $T_0 \in \mathcal{B}(T)$ *and* $T - T_0 \in \mathcal{B}(T)$*, and*

(iii) $g : T_0 \rightarrow \Re$ *is a continuous selection from* $G : T_0 \Rightarrow \Re$ *and* $C : T_0 \Rightarrow A$ *is upper hemicontinuous.*

Clearly, the upper semicontinuity of W in (A) is distinct from the Caratheodory property: one, neither implies, nor is implied by, the other. However, upper semicontinuity of W does imply that W is measurable. (B) is a generic continuity result. However, note that, while T_0 is "large" topologically, it is not necessarily so in other senses. For instance, consider the concrete model $T = [0, 1]$ with the Lebesgue measure. It is possible to construct $T_0 \subset [0, 1]$ with the properties listed above but with $\text{Leb}\,(T_0) = 0$, i.e., T_0 is negligible in the measure-theoretic sense.

Theorem 2.4.4. *Suppose* Γ *satisfies the assumptions of Lemma 2.3.1 (resp. 2.3.2).*

(A) If W *is Caratheodory, then there exists an optimal ECP for* Γ*.*

(B) If T *is compact metric,* W *is continuous and* $|C(t)| = 1$ *for every* t *in an open and dense subset of* T*, then there exists an optimal ECP for* Γ *that is continuous on* $T^0 \subset T$*, where* T^0 *has the same properties as* T_0*, as listed in Lemma 2.4.3(B).*

The following is an analogous result in the setting of Lemma 2.3.4 and Theorem 2.3.5.

Theorem 2.4.5. *If* Γ *satisfies the hypotheses of Lemma 2.3.4 and* W *is Caratheodory, then there exists an optimal ECP for* Γ*.*

2.5 Analysis of the application

In this section, we provide assumptions about the primitive data Γ^* that enable the application of the results of Sections 2.3 and 2.4 to prove the existence of (optimal) equilibrium coordination plans for Γ^*. As is to be expected, the trade-off evident in Section 2.3 will be reflected here. Roughly, the trade-off is that weaker assumptions on the action spaces A_i have to be paid for by stronger continuity assumptions on payoff functions u_i

and stronger assumptions regarding the principal's signal space T. The first result identifies an environment, specified by Assumptions (a), (b), (c) and (e), and a set of contracts, specified by Assumption (d), that satisfy the assumptions of Lemma 2.3.1.

Before stating our first result regarding Γ^*, we develop some notation. In the following two results, $\{L_i \mid i \in I\}$ is a family of topological vector spaces. Define $L = \prod_{i \in I} L_i$; given the product topology, it is easily verified that L is a topological vector space. If each L_i is a barreled space, then so is L ([28], IV.4.3).[7] Given a topological vector space M, let $\mathcal{L}$ be the space of continuous linear functions $f : L \to M$. Given $f \in \mathcal{L}$ and $A \subset L$, let f_A denote the restriction of f to A. Let $\mathcal{L}_A = \{f_A \mid f \in \mathcal{L}\}$. Let $e : M^A \times A \to M$ be the evaluation function defined by $e(f, a) = f(a)$. We now state our first result.

Theorem 2.5.1. *Given Γ^*, if*

(a) I is countable, and

for every $i \in I$,

(b) A_i is a convex, compact and metrizable barrel in a barreled space L_i,

(c) O_i is a compact subset of a normed space M,

(d) $T_i = \{\alpha + f_A \mid \alpha \in M \ \wedge \ f \in \mathcal{L}\} \cap O_i^A$, and

(e) v_i^* is continuous and quasi-concave,

then there exists an ECP for Γ^*. Moreover, if W is Caratheodory, then there exists an optimal ECP for Γ^*.

The contracts specified in Assumption (d) may be interpreted as two-part tariffs. We now consider Γ^* with the intention of applying Lemma 2.3.4.

Theorem 2.5.2. *Given Γ^*, if*

(a) I is countable, and

for every $i \in I$,

(b) A_i is a convex compact subset of a separable Banach space L_i with $\operatorname{Int} A_i \neq \emptyset$,

(c) O_i is a convex, separable and metrizable subset of a topological vector space,

(d) $(T_i, \mathcal{T}_i)$ is a complete measurable space, with

$$T_i = \bigcap_{a_{-i} \in A_{-i}} \{f \in O_i^A \mid f(a_{-i}, .) \text{ is linear}\}$$

and

(e) v_i^* is continuous and quasi-concave,

then there exists an ECP for Γ^**. Moreover, if* W *is Caratheodory, then there exists an optimal ECP for* Γ^**.*

3. The organization design problem

3.1 Principal's Role (α)

In this section we offer some observations on the organization design problem described by Role (α) in Section 1.1. We take as exogenously given the data so designated in Section 1.3. Consider an organization design $d = \{((T_i^d, \mathcal{T}_i^d), s_i^d, A_i^d), \sigma^d \mid i \in I\}$. Let a^d be an optimal ECP given design d. Finding and implementing a^d characterizes the principal's Role (β).

Suppose D is the set of organization designs and an optimal ECP a^d exists for each $d \in D$. As the example of Section 3.2 will show, the specification of D depends on the particular design context and the "information technology" (e.g., the language, vocabulary and grammar constraining the principal's messages to the agents) and "outcome-generation technology" (e.g., the production technology of a firm) relevant to that context. Given D, the principal's problem of choosing a design amounts to the problem of finding $d \in D$ such that

$$\int_\Omega P(d\omega) w \circ \sigma^d(\omega, a^d \circ s(\omega)) \geq \int_\Omega P(d\omega) w \circ \sigma^{d'}(\omega, a^{d'} \circ s(\omega))$$

for every $d' \in D$.

3.2 Example: designing a firm

Consider a firm with I as the set of agents, whose production operations consist of running the set of activities J; let I and J be finite. We define a design $\{((T_i, \mathcal{T}_i), s_i, A_i), \sigma \mid i \in I\}$ for this firm.

For every agent and every activity, let $[0,1]$ be the set of feasible activity levels. Agent i's action space is

$$A_i = \bigcap_{j \in J_i} \left\{ a \in [0,1]^J \;\middle|\; a_j = 0 \quad \wedge \quad \sum_{r \in J} a_r = 1 \right\}$$

where $J_i \subset J$. If $a \in A_i$, then a_j is interpreted as being agent i's contribution to activity $j \in J$. If $j \in J_i$, then agent i cannot contribute to activity j. Thus, $J - J_i$ is agent i's span of control in this design.

Fix the mapping $\sigma^* : \Omega \times \Re^J \to O$, where O is the outcome space and $\sigma^*(\omega, x)$ is the outcome that results when the state is ω and x is the profile of activity levels at which the various activities are operated. σ^*, which represents the firm's technology, is exogenously given and fixed. Given the action spaces $\{A_i \mid i \in I\}$, the outcome mapping $\sigma : \Omega \times A \to O$ is given by $\sigma(\omega, a) = \sigma^*(\omega, \sum_{i \in I} a_i)$. Clearly, the firm's "outcome-generation technology" is specified by the set of activities J, the production function σ^*, and the activity aggregation hypothesis that the profile of aggregate activity levels is separable across activities, and for each activity, the levels are additive across agents.

Suppose the state space is $\Omega = [0, 1]^K$, where K is finite; for simplicity, let $T = \Omega$ and let s be the identity mapping on Ω, i.e., the principal is fully informed about the state. The state may consist of data such as the firm's technology, the prevailing conditions in various input and output markets, and the contractual relations between the firm and buyers, suppliers and workers. Given $K_i \subset K$, suppose agent i's information set is $T_i = [0, 1]^{K_i}$ and $s_i = \pi_{K_i}$, where π_{K_i} projects $[0, 1]^K$ on $[0, 1]^{K_i}$. Thus, K_i is the set of information variables whose outcomes are communicated to agent i. The family $\{K_i \mid i \in I\}$ describes the firm's management information system. Clearly, this restricts the communication possibilities as a given agent either observes a given information variable or not; it is not possible, for instance, for an agent to observe the sum of two information variables instead of the two variables. Note that, while each agent's "message space" is uncountable, the "dimension" of the message space is constrained to be finite; so, while the "vocabulary" is rich, the "grammar" restricts the kinds of messages that can be fashioned using the vocabulary.

Suppose the set of designs D is identified with the set $(2^J \times 2^K)^I$, i.e., we accept the above-mentioned technological and communication constraints as given. Then, the number of ways of assigning action spaces and information are finite. Thus, the number of designs is finite and the problem of choosing among them is quite straightforward.

3.3 Some examples

In this section, we show *via* examples that the extent of information-sharing implied by the optimal design varies with the underlying setting. Consequently, a general monotonicity property connecting the optimal design and the extent of information sharing does not hold. The following data is common to all three examples: $\Omega = T = \{\omega_1, \omega_2\}$, $\mathcal{F} = \mathcal{T} = 2^\Omega$, $P = (1/2, 1/2)$, $s : \Omega \to T$ is the identity mapping, $O = \{o_1, o_2, o_3, o_4\}$,

$\mathcal{O} = 2^O$, $I = \{1,2\}$, $A_1 = \{U,D\}$, $A_2 = \{L,R\}$, λ is given by $\lambda(\omega,a) = 1/8$ for every $(\omega,a) \in \Omega \times A$, and $\sigma(\omega_1,.) : A \to O$ and $\sigma(\omega_2,.) : A \to O$ are represented by

$$\begin{array}{cc} & \begin{array}{cc} L & R \end{array} \\ \begin{array}{c} U \\ D \end{array} & \begin{pmatrix} o_1 & o_2 \\ o_3 & o_4 \end{pmatrix} \end{array} \quad \text{and} \quad \begin{array}{cc} & \begin{array}{cc} L & R \end{array} \\ \begin{array}{c} U \\ D \end{array} & \begin{pmatrix} o_4 & o_3 \\ o_2 & o_1 \end{pmatrix} \end{array}$$

respectively. For instance, if the state is ω_1 and the action choices are (D,R), then the outcome is o_4. Given the purpose of these examples, it is natural to keep the assignment of action spaces unchanged, while varying only the information dissemination mappings.

Example 3.3.1. *Suppose the agents' payoffs are generated by the mapping*

$$v(o) = (v_1(o), v_2(o)) = \begin{cases} (1,1), & \textit{if } o = o_1 \\ (4,0), & \textit{if } o = o_2 \\ (0,4), & \textit{if } o = o_3 \\ (2,2), & \textit{if } o = o_4 \end{cases}$$

and $w : O \to \Re$ *by* $w(o) = v_1(o) + v_2(o)$.

Thus, the agents' payoffs in the two states, $v \circ \sigma(\omega_1,.)$ and $v \circ \sigma(\omega_2,.)$, are given in bimatrix form by

$$\begin{array}{cc} & \begin{array}{cc} L & R \end{array} \\ \begin{array}{c} U \\ D \end{array} & \begin{pmatrix} 1,1 & 4,0 \\ 0,4 & 2,2 \end{pmatrix} \end{array} \quad \text{and} \quad \begin{array}{cc} & \begin{array}{cc} L & R \end{array} \\ \begin{array}{c} U \\ D \end{array} & \begin{pmatrix} 2,2 & 0,4 \\ 4,0 & 1,1 \end{pmatrix} \end{array}$$

respectively; this is *not* to say that the agents are playing the above bimatrix games in the two different states. If both agents have full information, i.e., $T_1 = T_2 = \Omega$, and s_1 and s_2 are the identity mappings, then the optimal ECP is

$$a(\omega) = \begin{cases} (U,L), & \text{if } \omega = \omega_1 \\ (D,R), & \text{if } \omega = \omega_2 \end{cases} \tag{3.3.2}$$

Consequently, the payoff profile is $(1,1)$ in both states, resulting in the principal's payoff being 2 in both states. On the other hand, if the agents are given no information, say $T_1 = T_2 = \{\omega_1\}$ (so that s_1 and s_2 are constant mappings), then

$$a(\omega) = \begin{cases} (U,R), & \text{if } \omega = \omega_1 \\ (D,L), & \text{if } \omega = \omega_2 \end{cases} \tag{3.3.3}$$

is an optimal ECP, which results in the principal getting payoff 4 in both states. It is clear from the structure of payoffs that the principal cannot improve upon the latter information structure.

Example 3.3.4. *Suppose the agents' payoffs are generated by the mapping*

$$v(o) = (v_1(o), v_2(o)) = \begin{cases} (2,2), & \text{if } o = o_1 \\ (2,1), & \text{if } o = o_2 \\ (1,2), & \text{if } o = o_3 \\ (0,0), & \text{if } o = o_4 \end{cases}$$

and $w : O \to \Re$ *by* $w(o) = v_1(o) + v_2(o)$.

Thus, the agents' payoffs in states ω_1 and ω_2 are given in bimatrix form by

$$\begin{array}{c} \\ U \\ D \end{array}\!\!\begin{array}{c} \begin{array}{cc} L & R \end{array} \\ \begin{pmatrix} 2,2 & 2,1 \\ 1,2 & 0,0 \end{pmatrix} \end{array} \quad \text{and} \quad \begin{array}{c} \\ U \\ D \end{array}\!\!\begin{array}{c} \begin{array}{cc} L & R \end{array} \\ \begin{pmatrix} 0,0 & 1,2 \\ 2,1 & 2,2 \end{pmatrix} \end{array}$$

respectively. If both agents have full information, then the optimal ECP is given by (3.3.2), resulting in the principal's payoff being 4 in both states. On the other hand, if the agents are given no information, then neither recommendation (U, L), nor recommendation (D, R), will induce obedience by them. Thus, in this case, the principal's expected payoff from any ECP must be less than 4; e.g., (3.3.3) is an ECP for this situation. It is clear from the structure of payoffs that the principal cannot improve upon the former information structure.

Example 3.3.5. *Suppose the agents' payoffs are generated by the mapping*

$$v(o) = (v_1(o), v_2(o)) = \begin{cases} (4,4), & \text{if } o = o_1 \\ (7,2), & \text{if } o = o_2 \\ (0,6), & \text{if } o = o_3 \\ (4,4), & \text{if } o = o_4 \end{cases}$$

and $w : O \to \Re$ *by* $w(o) = v_1(o) + v_2(o)$.

Thus, the agents' payoffs in states ω_1 and ω_2 are given in bimatrix form by

$$\begin{array}{c} \\ U \\ D \end{array}\!\!\begin{array}{c} \begin{array}{cc} L & R \end{array} \\ \begin{pmatrix} 4,4 & 7,2 \\ 0,6 & 4,4 \end{pmatrix} \end{array} \quad \text{and} \quad \begin{array}{c} \\ U \\ D \end{array}\!\!\begin{array}{c} \begin{array}{cc} L & R \end{array} \\ \begin{pmatrix} 4,4 & 0,6 \\ 7,2 & 4,4 \end{pmatrix} \end{array}$$

respectively. It is easy to confirm that, either if both agents are fully informed about the state or neither is informed about the state, then the optimal ECP is given by (3.3.2). Consequently, the principal's payoff is 8 in both states. A third possible information structure is that agent 1 is fully informed about the state but agent 2 is given no information, i.e., $T_1 = \Omega$ and s_1 is the identity mapping, while $T_2 = \{\omega_1\}$. In this situation, (3.3.3) is an optimal ECP. Thus, the optimal information structure in this example is intermediate between full revelation of the state, as in Example 3.3.4, and no revelation of information, as in Example 3.3.1.

4. Concluding remarks

We have studied a model of an organization that generalizes the classical notion of a team. Using this model, we define the notion of "organization design" in terms of choosing a system of information dissemination within the organization and assigning actions spaces to the members of the organization. The principal, i.e., the designer of the organization, is required to choose an organization design whose outcomes are optimal from the principal's perspective. The natural way to solve the problem is to adopt a two-step procedure. First, given a particular design, find the best outcome mapping that satisfies various informational and incentive constraints. Secondly, compare the best outcome mappings derived in the first step to arrive at the best design.

This paper is devoted to providing conditions that ensure the solvability of the problem entailed by the first step. We provide three sets of sufficient conditions. We also provide an application in which the principal and each agent are linked by an incentive contract. The setting of the application is such that the privateness of individual contracts is unimportant, but the very nature of the contracts is crucial for the existence of (optimal) ECPs. For instance, in one version of the application, restricting the class of contracts to a subset of affine mappings (a generalization of "two-part tariffs") is sufficient for guaranteeing the existence of optimal ECPs. In addition, we provide conditions on the primitives of the model that ensure the satisfaction of the critical reduced-form assumptions that are made in the main theorems. These lemmas may be of independent interest.

The second step optimization involved in completing the principal's program is unfinished business because the class of objects over which one must optimize, i.e., organization designs, is generally unstructured and there does not seem to be a way to reduce it to some canonical setting. As we have argued in Section 3.2, the set of feasible designs is intimately linked to the particular context and only a piecemeal theory may be possible.

Appendix A

Proof of Lemma 1.3.5. Let $a, b \in \mathcal{E}$, with a being an *ex post* optimal ECP, and let $\mu \in \Delta(T \times \Omega)$ be the unique measure such that $\mu(E \times F) = \int_E P \circ s^{-1}(dt)\Lambda(t, F)$ for $E \in \mathcal{T}$ and $F \in \mathcal{F}$. By definition, $\mu(T \times .) = \int_T P \circ s^{-1}(dt)\Lambda(t, .) = P(.)$. Therefore,

$$\int_\Omega P(d\omega) w \circ \sigma(\omega, a \circ s(\omega)) = \int_\Omega \mu(T \times d\omega) w \circ \sigma(\omega, a \circ s(\omega)) = \int_{T \times \Omega} \mu(dt \times d\omega) w \circ \sigma(\omega, a \circ s(\omega))$$

By the non-Cartesian version of the Fubini-Stone theorem ([27], Exercise 6.2.3),

$$\begin{aligned}\int_{T \times \Omega} \mu(dt \times d\omega) w \circ \sigma(\omega, a \circ s(\omega)) &= \int_T P \circ s^{-1}(dt) \int_\Omega \Lambda(t, d\omega) w \circ \sigma(\omega, a \circ s(\omega)) \\ &= \int_T P \circ s^{-1}(dt) \int_\Omega \Lambda(t, d\omega) w \circ \sigma(\omega, a(t)) \\ &= \int_T P \circ s^{-1}(dt) W(t, a(t))\end{aligned}$$

where the second equality is a consequence of Λ being regular. By an analogous argument, $\int_\Omega P(d\omega) w \circ \sigma(\omega, b \circ s(\omega)) = \int_T P \circ s^{-1}(dt) W(t, b(t))$. As $W(t, a(t)) \geq W(t, b(t))$ for every $t \in T$, we have

$$\begin{aligned}\int_\Omega P(d\omega) w \circ \sigma(\omega, a \circ s(\omega)) &= \int_T P \circ s^{-1}(dt) W(t, a(t)) \\ &\geq \int_T P \circ s^{-1}(dt) W(t, b(t)) = \int_\Omega P(d\omega) w \circ \sigma(\omega, b \circ s(\omega))\end{aligned}$$

as required. ■

Proof of Lemma 2.2.1. The proof proceeds in a number of steps.

(*i*) As Ω is compact by (a), it follows that $\Re^\Omega = \mathcal{C}(\Omega)$ and the compact-open topology of $\Re^\Omega$ is metrized by the supremum norm ([10], XII.8.2). With this norm, $\mathcal{C}(\Omega)$ is a Banach space ([11], V.7.17), and therefore, a Baire space ([10], XIV.4.1).

(*ii*) Let rca(Ω) be the set of regular countably additive real-valued set functions on $(\Omega, \mathcal{B}(\Omega))$. Given the total variation norm, rca(Ω) is a Banach space ([11], III.7.4) with the closed unit sphere S. By the Riesz representation theorem ([11], IV.6.3), rca(Ω) is isometrically isomorphic to the conjugate space of $\mathcal{C}(\Omega)$. As Ω is metric by (a), we have $\Delta(\Omega) \subset S$ ([11], III.9.22).

(*iii*) As Ω is compact by (a), and v_i and σ are continuous by (b), we have $v_i \circ \sigma(., b) \in \mathcal{C}(\Omega)$ for every $b \in A$. For every $(t, a) \in T \times A$, $\Lambda_i(s_i(t), a, .) \in \Delta(\Omega) \subset S$. Thus,

$G : T \times A^2 \to S \times \mathcal{C}(\Omega)$ is well-defined by $G(t, a, b) = (\Lambda_i(s_i(t), a, .), v_i \circ \sigma(., b))$. Define $L : S \times \mathcal{C}(\Omega) \to \Re$ by $L(\mu, f) = \int_\Omega \mu(d\omega) f(\omega)$. As $u_i = L \circ G$, it suffices to show that G and L are continuous.

For the rest of this proof, equip rca(Ω) *with the* $\mathcal{C}(\Omega)$ *topology* ([11], V.3.2). Consequently, by (*ii*) and Alaoglu's theorem ([11], V.4.2), S is compact.

(*iv*) Consider a net $(t_\alpha, a_\alpha) \subset T \times A$ converging to (t, a). Define the net $(\lambda_\alpha) \subset \Delta(\Omega)$ by $\lambda_\alpha(.) = \Lambda_i(s_i(t_\alpha), a_\alpha, .)$ and let $\lambda(.) = \Lambda_i(s_i(t), a, .)$. Consider $E \in \mathcal{B}(\Omega)$ such that $\lambda(\bar{E} - \operatorname{Int} E) = 0$. Using (b) and (c), we have $\lambda_\alpha(E) = \Lambda_i(s_i(t_\alpha), a_\alpha, E) \to \Lambda_i(s_i(t), a, E) = \lambda(E)$. Thus, (λ_α) converges to λ ([26], II.6.1). It follows that G_1 is continuous. Step (*i*) and (b) imply that G_2 is continuous ([10], XII.3.1).

(*v*) It remains to show that L is continuous. By the definition of the $\mathcal{C}(\Omega)$ topology, $L(., f)$ is continuous for every $f \in \mathcal{C}(\Omega)$. Consider $\mu \in S$ and $f \in \mathcal{C}(\Omega)$. To show the continuity of $L(\mu, .)$ at f, consider a sequence $(f_n) \subset \mathcal{C}(\Omega)$ converging to f. As $f_n \to f$ uniformly, $f_n \to f$ pointwise. As μ is countably additive, Lebesgue's theorem implies $\lim_{n\uparrow\infty} L(\mu, f_n) = L(\mu, f)$. Thus, $L(\mu, .)$ is continuous for every $\mu \in S$. Therefore, the identity $L(\mu, f) - L(\mu_0, f_0) = L(\mu - \mu_0, f - f_0) + L(\mu - \mu_0, f_0) + L(\mu_0, f - f_0)$ implies that it is sufficient to show that L is continuous at $(0, 0) \in S \times \mathcal{C}(\Omega)$.

Given $f \in \mathcal{C}(\Omega)$, as S is compact and $L(., f)$ is continuous, $L(S, f) \subset \Re$ is compact, and therefore, bounded. It follows that

$$\bigcup_{\mu \in S} L(\mu, F) = \{L(\mu, f) \mid \mu \in S \ \wedge \ f \in F\} = \bigcup_{f \in F} L(S, f) \subset \Re$$

is bounded for every finite set $F \subset \mathcal{C}(\Omega)$. Therefore, $\{L(\mu, .) \mid \mu \in S\}$ is a bounded subset of $\mathcal{L}(\mathcal{C}(\Omega), \Re)$ in the topology of pointwise convergence ([28], III.3.3). From this, and the fact that $\mathcal{C}(\Omega)$ is a Baire space by (*i*), it follows that $\{L(\mu, .) \mid \mu \in S\}$ is an equicontinuous set of mappings ([28], III.4.2). So, given an open neighborhood W of $0 \in \Re$, there exists an open neighborhood V of $0 \in \mathcal{C}(\Omega)$ such that $L(\mu, V) \subset W$ for every $\mu \in S$, i.e., $L(S \times V) \subset W$, as required. ■

Proof of Lemma 2.2.2. Given $(a, b) \in A^2$, the measurability of $u_i(., a, b)$ follows from the non-Cartesian version of the Fubini-Stone theorem ([27], Exercise 6.2.3) and the fact that s_i is measurable.

Given $t \in T$, $u_i(t, ., .)$ is continuous by an argument that is identical to that of Lemma 2.2.1, *modulo* trivial modifications as t is fixed in this argument. ■

Proof of Lemma 2.2.3. Given $(a,b) \in A^2$, the measurability of $u_i(.,a,b)$ follows from the non-Cartesian version of the Fubini-Stone theorem ([27], Exercise 6.2.3) and the fact that s_i is measurable.

We now fix $t \in T$ and show the continuity of $u_i(t,.,.)$.

(i) Equipped with the supremum norm, $B(\Omega,\mathcal{F})$ is a Banach space ([11], IV.5), and therefore, a Baire space ([10], XIV.4.1).

(ii) Let $\mathrm{ba}(\Omega)$ be the set of bounded additive real-valued set functions on $(\Omega,\mathcal{F})$. Given the total variation norm, $\mathrm{ba}(\Omega)$ is a Banach space ([11], III.7). Let S be the closed unit sphere of $\mathrm{ba}(\Omega)$. $\mathrm{ba}(\Omega)$ is the conjugate space of $B(\Omega,\mathcal{F})$ ([11], IV.5.1). Given $\mu \in \mathrm{ba}(\Omega)$ and $f \in B(\Omega,\mathcal{F})$, let $\langle \mu, f\rangle = \int_\Omega \mu(d\omega)f(\omega)$.

Henceforth, we equip $\mathrm{ba}(\Omega)$ *with the* $B(\Omega,\mathcal{F})$ *topology* ([11], V.3.2). By Alaoglu's theorem ([11], V.4.2), S is compact.

(iii) By definition, $\Lambda_i(s_i(t),a,.) \in \Delta(\Omega) \subset S$ for every $a \in A$. This and (c) imply that $G : A^2 \to S \times B(\Omega,\mathcal{F})$ is well-defined by $G(a,b) = (\Lambda_i(s_i(t),a,.), v_i \circ \sigma(.,b))$. Define $L : S \times B(\Omega,\mathcal{F}) \to \Re$ by $L(\mu,f) = \langle \mu, f\rangle$. Clearly, $u_i(t,.,.) = L \circ G(.,.)$. Thus, it suffices to show that G and L are continuous.

(iv) Continuity of G_2 follows directly from (c). We show that G_1 is continuous. Consider a net $(a_\alpha) \subset A$ converging to a. Define the net (λ_α) by $\lambda_\alpha(.) = \Lambda_i(s_i(t),a_\alpha,.)$ and let $\lambda(.) = \Lambda_i(s_i(t),a,.)$. We show that (λ_α) converges to λ. Given $E \in \mathcal{F}$, (d) implies

$$\lim_\alpha \langle \lambda_\alpha, 1_E\rangle = \lim_\alpha \Lambda_i(s_i(t),a_\alpha,E) = \Lambda_i(s_i(t),a,E) = \langle \lambda, 1_E\rangle$$

Bilinearity of $\langle .,.\rangle$ implies that an analogous argument holds when 1_E is replaced by a finite linear combination of functions drawn from $\{1_E \mid E \in \mathcal{F}\}$. Consider $f \in B(\Omega,\mathcal{F})$. By definition, there exists a sequence of functions (f_m) converging uniformly to f where each f_m is a finite linear combination of functions drawn from $\{1_E \mid E \in \mathcal{F}\}$. Then

$$\begin{aligned}\lim_\alpha \langle \lambda_\alpha, f\rangle = \lim_\alpha \langle \lambda_\alpha, \lim_m f_m\rangle &= \lim_\alpha \lim_m \langle \lambda_\alpha, f_m\rangle \\ &= \lim_m \lim_\alpha \langle \lambda_\alpha, f_m\rangle = \lim_m \langle \lambda, f_m\rangle = \langle \lambda, \lim_m f_m\rangle = \langle \lambda, f\rangle\end{aligned}$$

The first and last equalities follow from the definition of f. As λ_α and λ are countably additive, the second and fifth equalities follow from Lebesgue's theorem. The third equality is justified by the uniform convergence of (f_m) to f ([11], I.7.6). The fourth equality follows from the preceding argument as f_m is a finite linear combination of indicator functions.

(v) It remains to show that L is continuous. By the definition of the $\mathcal{B}(\Omega,\mathcal{F})$ topology, $L(.,f)$ is continuous for every $f \in \mathcal{B}(\Omega,\mathcal{F})$. Consider $\mu \in S$ and $f \in \mathcal{B}(\Omega,\mathcal{F})$. To show the continuity of $L(\mu, .)$ at f, consider a sequence $(f_n) \subset \mathcal{B}(\Omega, \mathcal{F})$ converging to f. As $f_n \to f$ uniformly, $f_n \to f$ pointwise. As μ is countably additive, Lebesgue's theorem implies $\lim_{n\uparrow\infty} L(\mu, f_n) = L(\mu, f)$. Thus, $L(\mu, .)$ is continuous for every $\mu \in S$. Therefore, the identity $L(\mu, f) - L(\mu_0, f_0) = L(\mu - \mu_0, f - f_0) + L(\mu - \mu_0, f_0) + L(\mu_0, f - f_0)$ implies that it is sufficient to show that L is continuous at $(0, 0) \in S \times \mathcal{B}(\Omega, \mathcal{F})$.

Given $f \in \mathcal{B}(\Omega, \mathcal{F})$, as S is compact and $L(., f)$ is continuous, $L(S, f) \subset \Re$ is compact, and therefore, bounded. It follows that

$$\bigcup_{\mu \in S} L(\mu, F) = \{L(\mu, f) \mid \mu \in S \ \wedge \ f \in F\} = \bigcup_{f \in F} L(S, f) \subset \Re$$

is bounded for every finite set $F \subset \mathcal{B}(\Omega, \mathcal{F})$. Therefore, $\{L(\mu, .) \mid \mu \in S\}$ is a bounded subset of $\mathcal{L}(\mathcal{B}(\Omega, \mathcal{F}), \Re)$ in the topology of pointwise convergence ([28], III.3.3). It follows from this and (i) that $\{L(\mu, .) \mid \mu \in S\}$ is an equicontinuous set of mappings ([28], III.4.2). So, given an open neighborhood W of $0 \in \Re$, there exists an open neighborhood V of $0 \in \mathcal{B}(\Omega, \mathcal{F})$ such that $L(\mu, V) \subset W$ for every $\mu \in S$, i.e., $L(S \times V) \subset W$, as required. ■

Proof of Corollary 2.2.4. Clearly, hypotheses (a), (b) and (d) are satisfied. We show that (c) is satisfied as well. (b') implies that $v_i \circ \sigma : \Omega \times A \to \Re$ is continuous. Consequently, $a \mapsto v_i \circ \sigma(., a)$ is a continuous mapping from A into $\Re^\Omega$ ([10], XII.3.1), where $\Re^\Omega$ is given the compact-open topology. As Ω is compact, $\Re^\Omega = \mathcal{C}(\Omega) \subset B(\Omega, \mathcal{B}(\Omega))$ and the compact-open topology on $\Re^\Omega$ is metrized by the supremum norm ([10], XII.8.2). ■

Proof of Lemma 2.2.6. Given $a \in A$, the measurability of $W(., a)$ follows from the non-Cartesian version of the Fubini-Stone theorem ([27], Exercise 6.2.3).

Fix $t \in T$ and consider a net $(a_\alpha) \subset A$ converging to a. Then $w \circ \sigma(., a_\alpha) \to w \circ \sigma(., a)$ pointwise. By Lebesgue's theorem, $W(t, a_\alpha) \to W(t, a)$. ■

Proof of Lemma 2.3.1. (A) Assumptions (c) and (d), and Theorems VI.3.1 and VI.3.2 in [4] imply that B_i is upper hemicontinuous, with nonempty and compact values. (d) also implies that B_i has convex values. Assumption (c) implies that A is convex, compact and metric. It follows that B is upper hemicontinuous ([12], Lemma A.4), with nonempty, convex and compact values.

(B) As the properties listed in (A) also hold for $B(t, .) : A \rightrightarrows A$ for every $t \in T$, Ξ has nonempty values by Theorem 4 in [6]. Define $f : T \times A \to T \times \operatorname{diag} A^2$ by $f(t, a) = (t, a, a)$. As A is metrizable, it is Hausdorff, and therefore, $\operatorname{diag} A^2$ is closed in A^2. (A) implies that $\operatorname{Gr} B$ is closed in $T \times A^2$. Thus, $\operatorname{Gr} B \cap (T \times \operatorname{diag} A^2)$ is closed in $T \times A^2$. As f is continuous,

$\mathrm{Gr}\,\Xi = f^{-1}(\mathrm{Gr}\,B \cap (T \times \mathrm{diag}\,A^2))$ is closed in $T \times A$. Thus, Ξ has closed values. As A is compact, Ξ has compact values.

(C) Let π_1 project $T \times A$ on T. Let E be closed in A. As T and A are compact, E and $T \times E$ are compact; moreover, (B) implies $\mathrm{Gr}\,\Xi$ is compact. Therefore, $\mathrm{Gr}\,\Xi \cap (T \times E)$ is compact. As π_1 is continuous, $\Xi^-(E) = \pi_1(\mathrm{Gr}\,\Xi \cap (T \times E))$ is compact. Since T is Hausdorff, $\Xi^-(E)$ is closed in T. Therefore, $\Xi^-(E) \in \mathcal{B}(T)$. Consequently, Ξ is measurable. Thus, Ξ is weakly measurable ([15], Theorem 3.5(i)).

(D) follows from the definitions. ■

Proof of Lemma 2.3.2. (A), (B) and (D) are proved as in Lemma 2.3.1.

(C) Define $\pi_2 : T \times A \to A$ by $\pi_2(t,a) = a$ and $\psi : T \times A \Rightarrow A$ by $\psi(t,a) = B(t,a) \cap \pi_2(t,a)$. Note that $\mathrm{Gr}\,\Xi = \{(t,a) \in T \times A \mid a \in B(t,a)\} = \{(t,a) \in T \times A \mid \psi(t,a) \neq \emptyset\} = \psi^-(A)$. (A) implies that $\mathrm{Gr}\,B$ is closed in $T \times A^2$. Using (b') and Lemma B.2, $\mathrm{Gr}\,B \in \mathcal{B}(T \times A^2) = \mathcal{B}(T) \times \mathcal{B}(A^2)$. Thus, B is weakly measurable ([15], Theorem 3.5(iii)). As π_2 is continuous, it is weakly measurable. Thus, as π_2 and B have closed values, ψ is measurable ([15], Theorem 4.1) and $\mathrm{Gr}\,\Xi = \psi^-(A) \in \mathcal{B}(T \times A) = \mathcal{B}(T) \times \mathcal{B}(A)$. So, (C) follows from Theorem 3.5(iii) in [15]. ■

Proof of Theorem 2.3.3. This follows from Lemma 2.3.1 (resp. 2.3.2) and Theorem 5.1 in [15]. ■

Proof of Lemma 2.3.4. In order to focus on the main steps, the proofs of some of the following arguments are collected in Appendix B. For steps (1) to (3), fix $i \in I$.

(1) Define $U_i : T \times A \times A_i \to \Re$ by $U_i(t,a,b) = u_i(t,a,a_{-i},b) - u_i(t,a,a)$ and $F_i : T \times A \Rightarrow A_i$ by $F_i(t,a) = \{b \in A_i \mid U_i(t,a,b) > 0\}$. Let $D_i = \{(t,a) \in T \times A \mid F_i(t,a) \neq \emptyset\}$. For every $t \in T$, let $D_i^t = \{a \in A \mid (t,a) \in D_i\}$, and for every $a \in A$, let $D_i^a = \{t \in T \mid (t,a) \in D_i\}$.

(2) By Lemma B.4, $F_i(t,.)$ is lower hemicontinuous for every $t \in T$. By Lemma B.5, F_i has open convex values in A_i, $\mathrm{Int}\,F_i(t,a) \neq \emptyset$ for every $(t,a) \in D_i$, and F_i is weakly measurable. By Lemma B.6, $D_i \in \mathcal{T} \times \mathcal{B}(A)$; D_i^t is open in A for every $t \in T$; and, for every $a \in A$, $D_i^a \in \mathcal{T}$.

(3) It follows from (2) and Theorem 3.2 in [17] that there exists $f_i : D_i \to A_i$ such that (i) $f_i(t,a) \in F_i(t,a)$ for every $(t,a) \in D_i$; (ii) $f_i(t,.) : D_i^t \to A_i$ is continuous for every $t \in T$; and (iii) $f_i(.,a) : D_i^a \to A_i$ is $(\mathcal{T} \cap D_i^a)/\mathcal{B}(A_i)$ measurable for every $a \in A$. Define $\phi_i : T \times A \Rightarrow A_i$ by

$$\phi_i(t,a) = \begin{cases} \{f_i(t,a)\}, & \text{if } (t,a) \in D_i \\ A_i, & \text{if } (t,a) \in (T \times A) - D_i \end{cases}$$

Clearly, ϕ_i has nonempty, convex and compact values. Define $\phi : T \times A \Rightarrow A$ by $\phi(t, a) = \prod_{i \in I} \phi_i(t, a)$. We now confirm our claims.

(A) These properties are inherited from the mappings $\{\phi_i \mid i \in I\}$.

(B) Let E be open in A_i. If $E = \emptyset$, then $\phi_i^-(E) = \emptyset \in \mathcal{T} \times \mathcal{B}(A)$. Suppose $E \neq \emptyset$. Then $\phi_i^-(E) = [(T \times A) - D_i] \cup f_i^{-1}(E)$. As $D_i \in \mathcal{T} \times \mathcal{B}(A)$ by (2), we have $(T \times A) - D_i \in \mathcal{T} \times \mathcal{B}(A)$. f_i is $[(\mathcal{T} \times \mathcal{B}(A)) \cap D_i]/\mathcal{B}(A_i)$ measurable ([17], Lemma 4.12). Therefore, $f_i^{-1}(E) = C \cap D_i$ for some $C \in \mathcal{T} \times \mathcal{B}(A)$. As $D_i \in \mathcal{T} \times \mathcal{B}(A)$ by (2), it follows that $f_i^{-1}(E) \in \mathcal{T} \times \mathcal{B}(A)$, and therefore, $\phi_i^-(E) \in \mathcal{T} \times \mathcal{B}(A)$.

Suppose E belongs to the basis for A, i.e., $E = \prod_{i \in I} E_i$, where E_i is open in A_i for every $i \in I$. Using (a), we have $\phi^-(E) = \cap_{i \in I} \phi_i^-(E_i) \in \mathcal{T} \times \mathcal{B}(A)$.

Let E be open in A. As A is second-countable, $E = \cup_{j \in J} E_j$ for a countable family $\{E_j \mid j \in J\}$ drawn from the basis for A. Therefore, $\phi^-(E) = \cup_{j \in J} \phi^-(E_j) \in \mathcal{T} \times \mathcal{B}(A)$.

As ϕ is weakly measurable, ϕ is measurable ([15], Theorem 3.5(ii)).

(C) Fix $t \in T$ and $i \in I$. Consider E open in A_i. If $E = A_i$, then $\phi_i(t, .)^+(E) = A$ is open in A. If $E \neq A_i$, then $\phi_i(t, .)^+(E) = f_i(t, .)^{-1}(E)$. As $f_i(t, .)$ is continuous by (3), $f_i(t, .)^{-1}(E) = D_i^t \cap U$ for some U open in A. As D_i^t is open in A by (2), it follows that $\phi_i(t, .)^+(E) = D_i^t \cap U$ is open in A. Thus, $\phi_i(t, .) : A \Rightarrow A_i$ is upper hemicontinuous. Using (a), $\phi(t, .) : A \Rightarrow A$ is upper hemicontinuous ([12], Lemma A.4).

(D) Suppose $a : T \to A$ is a measurable function such that $a(t) \in \phi(t, a(t))$ for every $t \in T$. Fix $t \in T$ and $i \in I$. If $(t, a(t)) \in D_i$, then $\phi_i(t, a(t)) = \{f_i(t, a(t))\}$. As $a_i(t) \in \phi_i(t, a(t))$, this implies $a_i(t) = f_i(t, a(t)) \in F_i(t, a(t))$. Thus, $U_i(t, a(t), a_i(t)) > 0$, which is a contradiction. So, $(t, a(t)) \in (T \times A) - D_i$. This implies $F_i(t, a(t)) = \emptyset$, i.e., $U_i(t, a(t), b) \leq 0$ for every $b \in A_i$. Consequently, a is an ECP.

Conversely, suppose $a : T \to A$ is an ECP. By definition, a is measurable. Fix $t \in T$ and $i \in I$. By definition, $U_i(t, a(t), b) \leq 0$ for every $b \in A_i$. Thus, $F_i(t, a(t)) = \emptyset$. Consequently, $(t, a(t)) \in (T \times A) - D_i$ and $a_i(t) \in A_i = \phi_i(t, a(t))$. Thus, $a(t) \in \phi(t, a(t))$ for every $t \in T$.

(E) Theorem 4 in [6] implies that Φ has nonempty values. We show that $A - \Phi(t)$ is open in A. Let $a \in A - \Phi(t)$. It follows that $a \in A - \phi(t, a)$. As A is Hausdorff and ϕ has compact values, ϕ has closed values. As A is compact Hausdorff, it is normal, and therefore, regular. Consequently, there exist open neighborhoods U_1 of a and U_2 of $\phi(t, a)$, such that $U_1 \cap U_2 = \emptyset$. As $\phi(t, .)$ is upper hemicontinuous, $\phi(t, .)^+(U_2)$ is open in A. Set $V = U_1 \cap \phi(t, .)^+(U_2)$. Note that $a \in V$ and V is open in A. Moreover, if $y \in V$, then

$\phi(t,y) \subset U_2 \subset A - U_1 \subset A - V$. Therefore, $y \in A - \phi(t,y)$, i.e., $y \in A - \Phi(t)$. Thus, $V \subset A - \Phi(t)$.

(F) Define $\pi_2 : T \times A \to A$ by $\pi_2(t,a) = a$. If E is open in A, then $\pi_2^{-1}(E) = T \times E \in \mathcal{T} \times \mathcal{B}(A)$. Therefore, π_2 is weakly measurable. By (B), ϕ is weakly measurable. Therefore, $\psi : T \times A \Rightarrow A$, defined by $\psi(t,a) = \phi(t,a) \cap \pi_2(t,a) = \phi(t,a) \cap \{a\}$, is measurable ([15], Theorem 4.1). Thus, $\mathrm{Gr}\,\Phi = \{(t,a) \in T \times A \mid a \in \phi(t,a)\} = \{(t,a) \in T \times A \mid \psi(t,a) \neq \emptyset\} = \psi^{-}(A) \in \mathcal{T} \times \mathcal{B}(A)$. As $(T, \mathcal{T})$ is complete, Φ is measurable ([15], Theorem 3.5(iii)).

(G) follows from (D) and the definition of Φ. ■

Proof of Theorem 2.3.5. This follows from Lemma 2.3.4 and Theorem 5.1 in [15]. ■

Proof of Lemma 2.4.1. (A) By Lemma 2.3.1, Ξ is measurable and has nonempty compact values. Therefore, G is measurable ([15], Theorems 6.5 and 3.5(ii)), with nonempty compact values. Therefore, $g(t) \in G(t)$ for every $t \in T$. Moreover, g is measurable ([15], Theorems 3.5(i) and 6.6).

(B) follows from Theorem 7.1 in [15]. ■

Proof of Lemma 2.4.2. The proofs of claims (A) and (B) follow the lines of the proof of Lemma 2.4.1. We show the claimed properties of C. Consider $t \in T$. (B) implies that $C(t) \neq \emptyset$. As $W(t,.)$ is continuous, $W(t,.)^{-1}([g(t),\infty))$ is closed in A. As $\Xi(t)$ is compact, $C(t)$ is closed.

As g is measurable and W is Caratheodory, $W - g$ is Caratheodory. Thus, the mapping $t \Rightarrow W(t,.)^{-1}([g(t),\infty)) = \{a \in A \mid W(t,a) - g(t) \geq 0\}$ is measurable ([15], Theorems 6.4). By Lemma 2.3.1, Ξ is measurable. Therefore, C is measurable ([15], Theorem 4.1). ■

Proof of Lemma 2.4.3. (A) By Lemmas 2.3.1 (resp. 2.3.2), Ξ has nonempty compact values and $\mathrm{Gr}\,\Xi$ is closed in $T \times A$. As A is compact Hausdorff, this means Ξ is upper hemicontinuous. Thus, $t \mapsto g(t) = \sup W(\{t\} \times \Xi(t)) = \sup\{W(t,a) \mid a \in \Xi(t)\}$ is upper semicontinuous ([4], Theorem VI.3.2).

Consider the set $[c,\infty)$ for $c \in \Re$. As g is upper semicontinuous, $g^{-1}([c,\infty))$ is closed in T. Consequently, $g^{-1}([c,\infty)) \in \mathcal{B}(T)$. This implies g is measurable as the family $\{[c,\infty) \mid c \in \Re\}$ generates $\mathcal{B}(\Re)$.

(B) As T and A are compact metric, they are complete and separable. As Ξ is upper hemicontinuous, there exists $T_0 \subset T$ such that (i) is satisfied and $\Xi : T_0 \Rightarrow A$ is continuous ([3], Theorem 1.4.13). (iii) follows immediately ([4], Theorems VI.3.1 and VI.3.2). We now confirm (ii).

It follows immediately from (i) that $T_0 \in \mathcal{B}(T)$ and $T - T_0 \in \mathcal{B}(T)$. As T is compact metric, T is complete, and therefore a Baire space ([10], XIV.4.1). T is of second category as it is a nonempty Baire space ([10], XI.10.5). (i) implies that T_0 is topologically complete ([10], XIV.8.3), and therefore, a Baire space ([10], XIV.4.1). (i) also implies that T_0 is dense in T ([10], XI.10.1). As each T_n is open and dense in T, $T - T_n$ is closed in T and $\text{Int}\,(T - T_n) = \emptyset$, i.e., each $T - T_n$ is nowhere dense. As $T - T_0 = T - \cap_{n\in\mathcal{N}} T_n = \cup_{n\in\mathcal{N}}(T - T_n)$, $T - T_0$ is a set of first category. If T_0 also is of first category, then T is of first category, a contradiction. So, T_0 is of second category. ■

Proof of Theorem 2.4.4. (A) By Lemma 2.4.1 (resp. 2.4.2), there exists measurable selection a from Ξ such that $g(t) = W(t, a(t))$ for every $t \in T$. By Lemma 2.3.1 (resp. 2.3.2), a is an ECP for Γ. Let b be an ECP for Γ. By Lemma 2.3.1 (resp. 2.3.2), b is a selection from Ξ. Thus, $W(t, b(t)) \in G(t)$ for every $t \in T$. Consequently, $W(t, b(t)) \leq \sup G(t) = g(t) = W(t, a(t))$ for every $t \in T$. The result follows from Lemma 1.3.5.

(B) Using (A), there exists an optimal ECP. Clearly, this ECP is a selection from C. Suppose T^* is an open and dense subset of T such that $|C(t)| = 1$ for every $t \in T^*$. Copying the argument of Lemma 2.4.3(B), $T^0 = T^* \cap T_0$ has all the claimed properties. Using Lemma 2.4.3(B), C is upper hemicontinuous on T^0. The result follows as C is singleton-valued on T^0. ■

Proof of Theorem 2.5.1. If the assumptions of Lemma 2.3.1 are satisfied, then the result follows from Theorems 2.3.3 and 2.4.5.

(a) and (b) imply that Assumptions (a) and (c) of Lemma 2.3.1 are satisfied. (1.4.1) implies $u_i(t, a, b) = v_i^* \circ t_i(b) = v_i^* \circ e(t_i, b)$. As (b) implies A is compact, e is continuous ([10], XII.2.4). Therefore, (e) implies that u_i is continuous. Assumption (d) of Lemma 2.3.1 is satisfied as (d) and (e) imply that $u_i(t, a, b_{-i}, .) = v_i^* \circ t_i(b_{-i}, .)$ is quasi-concave. It remains to verify that Assumption (b) of Lemma 2.3.1 is satisfied.

Suppose (1) T_i is closed in M^A, and (2) T_i is equicontinuous on A.

(c) implies that M^A is Hausdorff ([10], XII.1.3), and therefore, so is T_i. Let $T_i(a) = \{t_i(a) \in O_i \mid t_i \in T_i\}$ for $a \in A$. (c) implies that $\overline{T_i(a)}$ is compact for every $a \in A$. This fact, combined with (c) and (2), implies that $\overline{T_i}$ is compact ([10], XII.6.4). Therefore, (1) implies $T_i = \overline{T_i}$ is compact. As $T = \prod_{i\in I} T_i$, Assumption (b) of Lemma 2.3.1 is satisfied.

We note the following preliminary fact before verifying (1) and (2).

(0) Suppose $f \in M^A$ is linear on A, i.e., if $a, b \in A$, $\lambda, \mu \in \Re$ and $\lambda a + \mu b \in A$, then $f(\lambda a + \mu b) = \lambda f(a) + \mu f(b)$. We show that f has a unique extension $\bar{f} \in \mathcal{L}$, or equivalently, $f = \bar{f}_A$.

Consider $a \in L$. As A is a barrel, it is radial. Therefore, there exists $\lambda > 0$ such that $\lambda^{-1}a \in A$. Define $\bar{f}(a) = \lambda f(\lambda^{-1}a)$. Note that, if $\mu > 0$ is also such that $\mu^{-1}a \in A$, then $\lambda f(\lambda^{-1}a) = \lambda f(\mu^{-1}\mu\lambda^{-1}a) = \lambda\mu\lambda^{-1}f(\mu^{-1}a) = \mu f(\mu^{-1}a)$. Thus, the definition of $\bar{f}$ is independent of the choice of λ. We now check that $\bar{f}$ is linear. Consider $a, b \in L$. As A is radial, there exists $\lambda > 0$ such that $\lambda^{-1}a, \lambda^{-1}b, \lambda^{-1}(a+b) \in A$. Therefore, $\bar{f}(a+b) = \lambda f(\lambda^{-1}(a+b)) = \lambda f(\lambda^{-1}a + \lambda^{-1}b) = \lambda[f(\lambda^{-1}a) + f(\lambda^{-1}b)] = \lambda f(\lambda^{-1}a) + \lambda f(\lambda^{-1}b) = \bar{f}(a) + \bar{f}(b)$. Consider $a \in L$ and $\mu \in \Re$. As A is radial, there exists $\lambda > 0$ such that $\lambda^{-1}a, \lambda^{-1}\mu a \in A$. Therefore, $\bar{f}(\mu a) = \lambda f(\lambda^{-1}\mu a) = \lambda\mu f(\lambda^{-1}a) = \mu\bar{f}(a)$. It is straightforward to check that $\bar{f}$ is the unique linear extension of f. $\bar{f}$ is continuous as f is continuous at 0. Thus, f has a unique extension $\bar{f} \in \mathcal{L}$, or equivalently, $f = \bar{f}_A$. We now verify hypotheses (1) and (2).

(1) e is continuous ([10], XII.2.4). (c) implies that O_i is closed in M. Therefore, $O_i^A = \cap_{a \in A} e^{-1}(., a)(O_i)$ is closed in M^A. Consider a point of accumulation of T_i, say g. As g is also a point of accumulation of O_i^A, we have $g \in O_i^A$. It remains to show that $g = \alpha + \bar{f}_A$ for some $\alpha \in M$ and $\bar{f} \in \mathcal{L}$.

(b) and (c) imply that M^A is metrized by $d^+(f,g) = \sup\{\|f(a) - g(a)\| \mid a \in A\}$ ([10], XII.8.2). Consequently, there exists a sequence $(g_n) \subset T_i$ converging to g in the d^+ metric, and therefore also converging pointwise. Each g_n is of the form $g_n = \alpha_n + f_n$, where $\alpha_n \in M$ and $f_n \in \mathcal{L}(A)$. As A is a barrel, $0 \in A$. Define $\alpha = g(0)$ and $f = g - \alpha$. As g is continuous, so is f. As (g_n) converges to g, $\lim_{n\uparrow\infty} \|\alpha_n - \alpha\| = \lim_{n\uparrow\infty} \|g_n(0) - g(0)\| = 0$. For every $a \in A$, we have

$$\|f_n(a) - f(a)\| = \|g_n(a) - \alpha_n - g(a) + \alpha\| \leq \|g_n(a) - g(a)\| + \|\alpha_n - \alpha\|$$

It follows that $\lim_{n\uparrow\infty} \|f_n(a) - f(a)\| = 0$. Consider $a, b \in A$ and $\lambda, \mu \in \Re$ such that $\lambda a + \mu b \in A$. As each f_n is linear and M is a topological vector space, we have

$$f(\lambda a + \mu b) = \lim_{n\uparrow\infty} f_n(\lambda a + \mu b) = \lambda \lim_{n\uparrow\infty} f_n(a) + \mu \lim_{n\uparrow\infty} f_n(b) = \lambda f(a) + \mu f(b)$$

Thus, f is linear on A. By (0), $f = \bar{f}_A$ for some $\bar{f} \in \mathcal{L}$. Consequently, $g = \alpha + \bar{f}_A \in T_i$.

(2) We now show that T_i is equicontinuous on A. Let $B_r = \{x \in M \mid \|x\| < r\}$. As O_i is compact, it is bounded. Assume, without loss of generality, that $O_i \subset B_1$. If $\mathcal{L}_2 = \{f \in \mathcal{L} \mid f(A) \subset B_2\}$ is equicontinuous on L, then $\{\alpha + f \mid \alpha \in M \ \wedge \ f \in \mathcal{L}_2\}$ is equicontinuous on L, and therefore, on A. It follows that $\{\alpha + f_A \mid \alpha \in M \ \wedge \ f \in \mathcal{L}_2\}$ is equicontinuous on A. If $T_i \subset \{\alpha + f_A \mid \alpha \in M \ \wedge \ f \in \mathcal{L}_2\}$, then T_i is equicontinuous on A.

We first show that $T_i \subset \{\alpha + f_A \mid \alpha \in M \;\wedge\; f \in \mathcal{L}_2\}$. Consider $g = \alpha + f \in T_i$. Then, $\alpha = g(0) \in O_i \subset B_1$ and $f \in \mathcal{L}(A)$. For every $a \in A$, $\|f(a)\| = \|g(a) - \alpha\| \leq \|g(a)\| + \|\alpha\| < 2$, i.e., $f(A) \subset B_2$. By (0), there exists $\bar{f} \in \mathcal{L}$, such that $\bar{f}_A = f$. Thus, $g = \alpha + \bar{f}_A$. As $\bar{f}(A) = f(A) \subset B_2$, we have $\bar{f} \in \mathcal{L}_2$.

Finally, we show that $\mathcal{L}_2$ is equicontinuous. It suffices to show that $\mathcal{L}_2$ is bounded for the topology of pointwise convergence ([28], III.4.2). Consider $\mathcal{L}$ with the topology of pointwise convergence. We need to show that, for every open neighborhood $U \subset \mathcal{L}$ of $0 \in \mathcal{L}$, there exists $\lambda \in \Re$ such that $\mathcal{L}_2 \subset \lambda U$. It suffices to show this for every $U \subset \mathcal{L}$ that belongs to a local base at $0 \in \mathcal{L}$.

Given $a \in L$ and $E \subset M$, let $(a, E) = \{h \in \mathcal{L} \mid h(a) \in E\}$. A set U belonging to a local base for the topology of pointwise convergence on $\mathcal{L}$ at $0 \in \mathcal{L}$ is of the form $U = \cap_{j=1}^{k}(a^j, E)$ for some $\{a^1, \ldots, a^k\} \subset L$ and some open neighborhood E of $0 \in M$. For $\lambda \in \Re - \{0\}$, $\lambda U = \lambda \cap_{j=1}^{k} (a^j, E) = \cap_{j=1}^{k} \lambda(a^j, E)$.

As A is a barrel, $\mu_j = \sup\{\mu \in (0, 1] \mid \mu a^j \in A\}$ exists. Let $\nu > 0$ be such that $B_2 \subset \nu E$. Let $h \in \mathcal{L}_2$. Then, $h(A) \subset B_2 \subset \nu E$. Thus, $h(\mu_j a^j) \in \nu E$, i.e., $h(a^j) \in \lambda_j E_j$ where $\lambda_j = \nu/\mu_j$. Therefore, $h \in (a^j, \lambda_j E)$. Setting $\lambda = \max\{\lambda_j \mid j = 1, \ldots, k\}$, we have $h \in \cap_{j=1}^{k}(a^j, \lambda E) = \cap_{j=1}^{k} \lambda(a^j, E) = \lambda U$. Thus, $\mathcal{L}_2 \subset \lambda U$. ■

Proof of Theorem 2.5.2. If the assumptions of Lemma 2.3.4 are satisfied, then the result follows from Theorems 2.3.5 and 2.4.5. (a), (b) and (d) imply that Assumptions (a), (b) and (c) of Lemma 2.3.4 are satisfied. By (1.4.1), $u_i(t, a, b) = v_i^* \circ t_i(b) = v_i^* \circ e(\pi_i(t), b)$. π_i is continuous as it is the projection mapping. e is continuous as A is compact ([10], XII.2.4). Thus, u_i is continuous, and consequently, Caratheodory. (d) and (e) combine to imply that $u_i(t, a, b_{-i}, .)$ is quasi-concave. Thus, Assumption (d) of Lemma 2.3.4 is satisfied. ■

Appendix B

Lemma B.1. *Suppose*

(a) $(X, \mathcal{T})$ is a topological space, $Y \subset X$, and $\mathcal{T}_Y$ is the subspace topology on Y,

(b) $\mathcal{B}(X)$ (resp. $\mathcal{B}(Y)$) is the σ-algebra generated on X (resp. Y) by $\mathcal{T}$ (resp. $\mathcal{T}_Y$).

Then, $\mathcal{B}(Y) = \{E \cap Y \mid E \in \mathcal{B}(X)\}$.

Proof. Let $F \in \mathcal{T}_Y$. Then, there exists $E \in \mathcal{T}$ such that $F = E \cap Y$. As $E \in \mathcal{T}$, we have $E \in \mathcal{B}(X)$. Consequently, $\mathcal{T}_Y \subset \{E \cap Y \mid E \in \mathcal{B}(X)\}$. As $\{E \cap Y \mid E \in \mathcal{B}(X)\}$ is a σ-algebra, it follows that $\mathcal{B}(Y) \subset \{E \cap Y \mid E \in \mathcal{B}(X)\}$.

Conversely, let $\mathcal{C} = \{A \cup (B - Y) \mid A \in \mathcal{B}(Y) \ \wedge \ B \in \mathcal{B}(X)\}$. We show below that $\mathcal{B}(Y) = \{E \cap Y \mid E \in \mathcal{C}\}$. Therefore, we need to show that $\{E \cap Y \mid E \in \mathcal{B}(X)\} \subset \{E \cap Y \mid E \in \mathcal{C}\}$. For this, it is sufficient to show that $\mathcal{B}(X) \subset \mathcal{C}$. It is easy to show that $\mathcal{C}$ is a σ-algebra on X. Moreover, suppose $E \in \mathcal{T}$. Then, $E \cap Y \in \mathcal{T}_Y \subset \mathcal{B}(Y)$ and $E \in \mathcal{B}(X)$. It follows that $E = (E \cap Y) \cup (E - Y) \in \mathcal{C}$. Therefore, $\mathcal{T} \subset \mathcal{C}$. It follows that $\mathcal{B}(X) \subset \mathcal{C}$.

To complete the argument, we show that $\mathcal{B}(Y) = \{E \cap Y \mid E \in \mathcal{C}\}$. If $F \in \mathcal{B}(Y)$, then $F \subset Y$ and $F = F \cup \emptyset = F \cup (\emptyset - Y) \in \mathcal{C}$. Consequently, $\mathcal{B}(Y) \subset \{E \cap Y \mid E \in \mathcal{C}\}$. Conversely, let $A \in \mathcal{B}(Y)$ and $B \in \mathcal{B}(X)$. Then, $[A \cup (B - Y)] \cap Y = (A \cap Y) \cup [(B - Y) \cap Y] = A \cap Y = A \in \mathcal{B}(Y)$. Therefore, $\{E \cap Y \mid E \in \mathcal{C}\} \subset \mathcal{B}(Y)$. ■

Lemma B.2. *Suppose*

(a) $\{X_i \mid i \in I\}$ *is a countable family of separable metric spaces,*

(b) for every $i \in I$, X_i *is given the Borel* σ*-algebra* $\mathcal{B}(X_i)$*, and*

(c) $X = \prod_{i \in I} X_i$ *is given the product topology and the Borel* σ*-algebra* $\mathcal{B}(X)$*.*

Then, $\mathcal{B}(X)$ *is identical to the product* σ*-algebra generated on* X *by the family of* σ*-algebras* $\{\mathcal{B}(X_i) \mid i \in I\}$*.*

Proof. See Theorem I.1.10 in [26]. ■

Lemma B.3. *Given the assumptions of Lemma 2.3.4,*

(A) $U_i(t, ., .)$, $U_i(t, a, .)$ *and* $U_i(t, ., b)$ *are continuous for every* $t \in T$, $a \in A$ *and* $b \in A_i$*,*

(B) $U_i(., a, b) : T \to \Re$ *is measurable for every* $(a, b) \in A \times A_i$*, and*

(C) $U_i(., ., b) : T \times A \to \Re$ *is measurable for every* $b \in A_i$*.*

Proof. (A) and (B) follow immediately from the assumptions. (C) follows from (A), (B) and Theorem 6.1 in [15]. ■

Lemma B.4. *Given the assumptions of Lemma 2.3.4,* $F_i(t, .)$ *is lower hemicontinuous for every* $t \in T$*.*

Proof. Fix $t \in T$. Lemma B.3(A) implies $\operatorname{Gr} F_i(t, .) = \{(a, b) \in A \times A_i \mid b \in F_i(t, a)\} = \{(a, b) \in A \times A_i \mid U_i(t, a, b) > 0\}$ is open in $A \times A_i$. Let E be open in A_i. Note that

$$\begin{aligned} F_i(t, .)^-(E) &= \{a \in A \mid F_i(t, a) \cap E \neq \emptyset\} \\ &= \pi_1\left(\{(a, b) \in A \times A_i \mid b \in F_i(t, a)\} \cap A \times E\right) \\ &= \pi_1(\operatorname{Gr} F_i(t, .) \cap A \times E) \end{aligned}$$

where $\pi_1 : A \times A_i \to A$ is the projection $\pi_1(a, b) = a$. As $\operatorname{Gr} F_i(t, .)$ and $A \times E$ are open in $A \times A_i$, $\operatorname{Gr} F_i(t, .) \cap A \times E$ is open in $A \times A_i$. As π_1 is an open mapping, $F_i(t, .)^-(E)$ is open in A. ■

Lemma B.5. *Given the assumptions of Lemma 2.3.4,*

(A) F_i *is weakly measurable,*

(B) $F_i(t,a)$ *is convex and open in* A_i *for every* $(t,a) \in T \times A$*, and*

(C) $\operatorname{Int} F_i(t,a) \neq \emptyset$ *for every* $(t,a) \in D_i$*.*

Proof. (A) By definition, for every $E \subset A_i$, we have

$$\begin{aligned} F_i^-(E) &= \{(t,a) \in T \times A \mid F_i(t,a) \cap E \neq \emptyset\} \\ &= \bigcup_{b \in E} \{(t,a) \in T \times A \mid b \in F_i(t,a)\} = \bigcup_{b \in E} \{(t,a) \in T \times A \mid U_i(t,a,b) > 0\} \end{aligned}$$

As A_i is compact metric, it is separable. Let C be a countable set that is dense in A_i and let E be open in A_i. Then, Lemma B.3(A) implies

$$F_i^-(E) = \bigcup_{b \in E \cap C} \{(t,a) \in T \times A \mid U_i(t,a,b) > 0\}$$

By Lemma B.3(C), $\{(t,a) \in T \times A \mid U_i(t,a,b) > 0\} \in \mathcal{T} \times \mathcal{B}(A)$ for every $b \in E \cap C$. As $E \cap C$ is countable, $F_i^-(E) \in \mathcal{T} \times \mathcal{B}(A)$.

(B) follows from Assumption (d) and Lemma B.3(A).

(C) follows from (B) and the definition of D_i. ■

Lemma B.6. *Given the assumptions of Lemma 2.3.4,*

(A) $D_i \in \mathcal{T} \times \mathcal{B}(A)$*,*

(B) D_i^t *is open in* A *for every* $t \in T$*, and*

(C) $D_i^a \in \mathcal{T}$ *for every* $a \in A$*.*

Proof. (A) Lemma B.5 implies $D_i = F_i^-(A_i) \in \mathcal{T} \times \mathcal{B}(A)$.

(B) Fix $t \in T$. Then, $D_i^t = \{a \in A \mid F_i(t,a) \neq \emptyset\} = \cup_{b \in A_i} \{a \in A \mid U_i(t,a,b) > 0\}$. Lemma B.3(A) implies $\{a \in A \mid U_i(t,a,b) > 0\}$ is open in A for every $b \in A_i$.

(C) Fix $a \in A$. Lemma B.3(A) implies that

$$D_i^a = \{t \in T \mid F_i(t,a) \neq \emptyset\} = \bigcup_{b \in A_i} \{t \in T \mid U_i(t,a,b) > 0\} = \bigcup_{b \in C} \{t \in T \mid U_i(t,a,b) > 0\}$$

Lemma B.3(B) implies $\{t \in T \mid U_i(t,a,b) > 0\} \in \mathcal{T}$ for every $b \in C$. As C is countable, $D_i^a \in \mathcal{T}$. ■

Notes

1. By "economic" we mean a theory based on rational choice, as opposed to one that seeks other explanations, e.g., evolutionary theories.

2. A related question is: given an organization design, what features of the environment will rationalize that design? Answers to this question, say in the case of the firm, may be found in the literature following [1], [2], [8], [16], [20], [23] and [30]. If observed designs are interpreted as optimal responses to their environments, then a solution of the general normative problem also contains the answer to the above question.

3. As shown in [24], using an argument combining the "revelation principle" with an "obedience principle", this set-up of the problem is canonical, i.e., a large class of seemingly different principal-agent problems can be re-cast in this set-up without any substantive loss of generality.

4. The general model is more complicated in that the principal offers a mechanism in the first stage, which specifies message spaces for the principal and the agent and an outcome function that maps message profiles to outcomes. By the revelation principle, the mechanism offered can be restricted to be a direct mechanism that induces truthful revelation of types in equilibrium. This, along with the fact that the agent has no private information, implies that the types-to-outcome mapping resulting from the truthful implementation of the mechanism coincides with the outcome function itself, as stated above.

5. We simply assume the existence of a regular conditional distribution. However, existence of such a function is guaranteed if Ω and T are separable standard measurable spaces ([26], V.8.1).

6. This requirement can be dropped if the relevant linear spaces are finite dimensional.

7. Examples of barreled spaces include every l.c.s. that is a Baire space, in particular Banach spaces and Fréchet spaces ([28], II.7.1). See [28] for other terminology used in the proof of this result.

References

[1] Alchian, A. and H. Demsetz. 1972. Production, information costs and economic organization. *American Economic Review* 62, 777-795.

[2] Arrow, K. J. 1974. *The Limits of Organization.* W. W. Norton.

[3] Aubin, J-P. and H. Frankowska. 1990. *Set-valued Analysis.* Birkhäuser.

[4] Berge, C. 1963. *Topological Spaces.* Macmillan.

[5] Bernheim, D. and M. Whinston. 1986. Common agency. *Econometrica* 54, 923-942.

[6] Browder, F. E. 1968. The fixed point theory of multi-valued mappings in topological vector spaces. *Mathematische Annalen* 177, 283-301.

[7] Cho, I. K. and D. Kreps. 1987. Signaling games and stable equilibria. *Quarterly Journal of Economics* 102, 179-221.

[8] Coase, R. H. 1937. The nature of the firm. *Economica* 4, 386-405.

[9] Dixit, A., G. M. Grossman and E. Helpman. 1997. Common agency and coordination: general theory and application to government policy making. *Journal of Political Economy* 105, 752-769.

[10] Dugundji, J. 1989. *Topology.* Wm. C. Brown.

[11] Dunford, N. and J. T. Schwartz. 1988. *Linear Operators, Part I: General Theory*, Wiley Interscience.

[12] Fan, K. 1952. Fixed-point and minimax theorems in locally convex topological linear spaces. *Proceedings of the National Academy of Sciences* 38, 121-126.

[13] Fudenberg, D. and J. Tirole. 1991. *Game Theory.* MIT Press.

[14] Grossman, S. and O. Hart. 1983. An analysis of the principal-agent problem. *Econometrica* 51, 7-46.

[15] Himmelberg, C. J. 1975. Measurable relations. *Fundamenta Mathematicae* 87, 53-72.

[16] Holmstrom, B. 1982. Moral hazard in teams. *Bell Journal of Economics* 13, 324-340.

[17] Kim, T., K. Prikry and N. C. Yannelis. 1987. Caratheodory-type selections and random fixed point theorems. *Journal of Mathematical Analysis and Applications* 122, 393-407.

[18] Laffont, J.-J. and J. Tirole. 1986. Using cost observation to regulate firms. *Journal of Political Economy* 94, 614-641.

[19] Lazear, E. and S. Rosen. 1981. Rank-order tournaments as optimum labor contracts. *Journal of Political Economy* 89, 841-864.

[20] Marschak, J. and R. Radner. 1972. *The Theory of Teams.* Yale University Press.

[21] Maskin, E. and J. Tirole. 1990. The principal-agent relationship with an informed principal: the case of private values. *Econometrica* 58, 379-409.

[22] Maskin, E. and J. Tirole. 1992. The principal-agent relationship with an informed principal, II: common values. *Econometrica* 60, 1-42.

[23] Mirrlees, J. 1976. The optimal structure of incentives and authority within an organization. *Bell Journal of Economics* 7, 105-131.

[24] Myerson, R. B. 1982. Optimal coordination mechanisms in generalized principal-agent problems. *Journal of Mathematical Economics* 10, 67-81.

[25] Myerson, R. B. 1983. Mechanism design by an informed principal. *Econometrica* 51, 1767-1798.

[26] Parthasarathy, K. R. 1967. *Probability Measures on Metric Spaces.* Academic Press.

[27] Rao, M. M. 1987. *Measure Theory and Integration.* John Wiley & Sons.

[28] Schaefer, H. H. 1986. *Topological Vector Spaces.* Springer-Verlag.

[29] Shavell, S. 1979. Risk sharing and incentives in the principal and agent relationship. *Bell Journal of Economics* 10, 55-73.

[30] Williamson, O. E. 1967. Hierarchical control and optimum firm size. *Journal of Political Economy* 75, 123-138.

[31] Wilson, R. 1968. The theory of syndicates. *Econometrica* 36, 119-132.