On the existence of optimal contract mechanisms for incomplete information principal-agent models

Erik J. Balder
Mathematical Institute, University of Utrecht
Utrecht, the Netherlands

June 11, 1997

Two abstract results are given for the existence of optimal contract selection mechanisms in principal-agent models; by a suitable reformulation of the (almost) incentive compatibility constraint, they deal with both single- and multi-agent models. In particular, it is shown that the existence results in a series of papers by F.H. Page [9, 10, 12, 13] follow from these two results.

1 Introduction

In a series of papers, F.H. Page presented existence results for optimal contract selection mechanisms for a number of principal-agent models [9, 12, 13]. These models involve both moral hazard and incomplete information aspects (in the sense of e.g. Harsanyi [6] and Meyerson [8]). A quite related existence result was also given by Page for the existence of an optimal Stackelberg strategy in a game with incomplete information [10]. In some form or another, all these models deal with the maximization of the principal's expected utility function

$$\int_T U(t, f(t))\nu(dt)$$

over the set of all contract selection mechanisms $f : T \to K$ that are (i) individually rational and (ii) incentive compatible. Here ν stands for the principal's subjective probability distribution of the types, T being the set of all types that the agent(s) can report to have. By working with a mechanism f, the principal is committed to using the contract $f(t)$ from K if type t is reported by the agent(s).

The nature of the incentive compatibility constraint (see Definitions 2.1, 2.4), which is imposed so as to induce truthful reporting by the agent(s), causes the existence problem to be quite nonstandard, the constraint of *incentive compatibility* (used in [10]) presenting a technically more challenging problem than *almost sure incentive compatibility* [9, 12, 13]. The inherent technical difficulties were solved by Page, by adopting the use of an abstract Komlós theorem from [1]. This theorem has the important advantage of supplying almost everywhere *pointwise* convergence information about arithmetic averages in situations where traditionally only weak convergence information, much more macroscopic in nature, is available. In yet another development, C.M. Khan showed that a less refined analogue of the abstract Komlós theorem, viz. Mazur's theorem, can be used to deal with the existence question, provided that one weakens the notion of a.s. incentive compatibility considerably. This would seem to confirm a pattern signaled above: the more refined the incentive compatibility constraint, the more refined the analysis to solve the associated existence questions must be. Apart from this essential difference with Page's existence results, another conspicuous contrast is that Khan does not require the set of contracts to be compact. Instead, he imposes monotonicity and other structural conditions upon his utility function (his space of contracts being two-dimensional). However, as Khan himself shows, his incentive compatibility and related conditions force the problem to be equivalent to one with a compact set of contracts. In spite of its more explicit nature, [7] does not provide explicit conditions under which the expected utility functional is upper semi-continuous and existence actually occurs (see [7, p. 126]).

The purpose of the present paper is to produce a systematic, general approach to existence problems involving incentive compatibility. Two principal results are formulated; together they apply to all situations considered by Page in the papers mentioned above, and in other situations as well [3]. The essential steps of the approach all stem from the work by Page. This paper serves to provide a unifying platform for those ideas. Its use of misspecification subsets of types seems a new idea. This is used in the definition of (a.s.) incentive compatibility. Together with the use of a suitable σ-algebra on the space T of all types, it allows for the simultaneous treatment of single- and multi-agent models.

2 Principal existence results

The treatment in this section is at a rather high level of abstraction; of course, this is unavoidable if one wishes to bring many applications together. The reader is advised to refer continuously to section 3 for motivation. Proofs of all results in this section can be found in section 4.

Let $(T, \mathcal{T}, \nu)$ be a probability space; elements of T represent the various *types* (or *type-vectors*) of the agent(s). Let K be a subset of a Hausdorff locally convex topological space E; elements of K will be called *contracts*, and measurable functions from T into K will be called *contract selection mechanisms* for the principal. Once she/he adopts the contract selection mechanism $f : T \to K$, the principal is committed to offering the contract $f(t)$ upon being reported the type t. Here measurability is defined with respect to $\mathcal{T}$ and the Borel σ-algebra $\mathcal{B}(K)$ on K. As usual, the topological dual of E (i.e., the set of all linear continuous functionals on E) is denoted by E^*. For $x \in E$ and $x^* \in E^*$ we shall write $< x, x^* > := x^*(x)$.

Assumption 2.1 *The set K of contracts is convex compact and metrizable for the relative topology (as inherited from E).*

Let $\Gamma : T \to 2^K$ be a multifunction. For a type t agent, $\Gamma(t)$ represents the set of *rational* contracts for such an agent. No measurability conditions whatsoever will be needed for Γ.

Assumption 2.2 *The set $\Gamma(t) \subset K$ of rational contracts is convex and closed for each $t \in T$.*

Below we allow for two different notions of rationality and incentive compatibility of mechanisms, depending on whether they have to hold surely or just almost surely (with respect to the subjective probability ν). For each of these situations we shall present an existence result, respectively in Theorems 2.1 and 2.2. We now prepare for the first of these:

Definition 2.1 (a.s. individual rationality) *A contract selection mechanism f is said to be* almost surely individually rational *if*

$$f(t) \in \Gamma(t) \text{ for } \nu\text{-almost every } t \in T.$$

Recall that this states that there exists a subset N in $\mathcal{T}$, with $\nu(N) = 0$ (i.e., N is a *null set* for ν), such that $f(t) \in \Gamma(t)$ for all $t \in T \backslash N$. Observe that N, the "exceptional null set" for f's pointwise belonging to Γ, is allowed to vary with f. Although it is not strictly required, N might be the empty set in some instances (the stronger Definition 2.4 exactly calls for such emptiness). From now on we shall often write "almost every" or even "a.e." instead of the more formal "ν-almost every". By $\mathcal{S}_\Gamma^{as}$ we shall denote the set of a.s. individually rational mechanisms. It should be stressed that measurability is required for the mechanisms, but not for Γ. Let $V : T \times K \to \mathbf{R}$ be the *agent's utility function.* An agent of type t is supposed to use $V(t, \cdot)$ as his/her utility function.

Assumption 2.3 *For each $t \in T$ the type t agent's utility function $V(t, \cdot)$ is affine and continuous on the set $\Gamma(t)$ of their rational contracts.*

Note that a type t agent will not have an incentive to misreport under the mechanism f if what he/she gets under the contract $f(t)$ is at least as good as what he/she gets by misreporting his/her type as t'. Both these options are evaluated by such an agent via the utility function $V(t, \cdot)$; hence the first option is worth $V(t, f(t))$ utils to the agent, and the second one $V(t, f(t'))$ utils. The Definitions 2.2 and 2.5 state that (almost surely) the second utility amount will not exceed the first one. The following device allows us to treat multi- and single-agent models simultaneously. Let $F : T \to 2^T$ be a *misspecification* multifunction; for each $t \in T$, $F(t)$ is the set from which type t agents can opt to *misreport* to the principal. No form of measurability whatsoever is required for F. In several applications below it is natural to use $F(t) \equiv T$; however, when treating the multi-agent setup it is essential to use a different misspecification multifunction F.

Definition 2.2 (a.s. incentive compatibility) *A contract selection mechanism f is said to be* almost surely incentive compatible *if for almost every t*

$$V(t, f(t)) \geq V(t, f(t')) \text{ for all } t' \in F(t).$$

This notion is encountered in e.g. [9, 11, 12, 13]. In [7] Khan requires the above inequality to hold only for all t' not in some ν-null set. This would seem to be a weaker and intuitively less convincing notion. Let $\mathcal{S}_{IC}^{as}$ denote the set of all mechanisms that are a.s. incentive compatible.

Let $U : T \times K \to [-\infty, +\infty)$ be the *principal's utility function.* The assumptions to follow allow us to define the principal's *expected utility* function. Namely, the expected utility of the mechanism $f \in \mathcal{S}_\Gamma$ is given by

$$I_U(f) := \int_T U(t, f(t))\nu(dt),$$

where ν gives the principal's (subjective) probability distribution of the various types.

Assumption 2.4 *The principal's utility function U is $\mathcal{T} \times \mathcal{B}(K)$-measurable.*

Let us note that instead of product measurability one could do with mere $\mathcal{T}$-measurability of $U(\cdot, f(\cdot))$ for all $f \in \mathcal{S}_\Gamma^{as} \cap \mathcal{S}_{IC}^{as}$ (in fact, it would even be possible to abandon measurability for the principal's utility altogether by adopting outer integration [1, 2]).

Assumption 2.5 *There exists a ν-integrable function $\psi : T \to \mathbf{R}$ such that for each $f \in \mathcal{S}_\Gamma^{as} \cap \mathcal{S}_{IC}$*

$$U(t, f(t)) \leq \psi(t) \ a.e.$$

Assumption 2.6 *For each type t in T the function $U(t, \cdot)$ is concave and upper semicontinuous on $\Gamma(t)$.*

Assumption 2.7 *The set $\mathcal{S}_\Gamma^{as} \cap \mathcal{S}_{IC}^{as}$ of all a.s. individually rational, a.s. incentive compatible mechanisms is nonempty.*

Definition 2.3 *The* principal's maximization problem (P^{as}) *consists of maximizing the expected utility $I_U(f)$ over all mechanisms f in $\mathcal{S}_\Gamma^{as} \cap \mathcal{S}_{IC}^{as}$.*

Theorem 2.1 *Under the Assumptions 2.1–2.7 there exists an optimal mechanism for the principal's utility maximization problem (P^{as}).*

Next, we prepare for a closely related but stronger existence result, for which also stronger conditions are needed.

Definition 2.4 (individual rationality) *A contract selection mechanism f is said to be* individually rational *if*

$$f(t) \in \Gamma(t) \text{ for every } t \in T.$$

By $\mathcal{S}_\Gamma$ we denote the set of individually rational mechanisms. Clearly, this set is contained in the earlier set $\mathcal{S}_\Gamma^{as}$. Note that this new notion of individual rationality no longer depends on the subjective probability distribution ν. A similar comment applies to the following notion, which ensures that none of the agents has an incentive to misreport her/his type.

Definition 2.5 (incentive compatibility) *A contract selection mechanism f is said to be* incentive compatible *if for all t*

$$V(t, f(t)) \geq V(t, f(t')) \text{ for all } t' \in F(t).$$

By $\mathcal{S}_{IC}$ we denote the set of all mechanisms that are incentive compatible. Clearly, $\mathcal{S}_{IC}$ is contained in the earlier set $\mathcal{S}_{IC}^{as}$.

Assumption 2.8 *The set $\mathcal{S}_\Gamma \cap \mathcal{S}_{IC}$ of all individually rational, incentive compatible mechanisms is nonempty.*

Evidently, this is a little stronger than Assumption 2.7. Moreover, an extra measurability condition will be used for V:

Assumption 2.9 *The utility function V is $\mathcal{T} \times \mathcal{B}(K)$-measurable.*

Definition 2.6 *The* principal's maximization problem (P) *consists of maximizing the expected utility $I_U(f)$ over all mechanisms f in $\mathcal{S}_\Gamma \cap \mathcal{S}_{IC}$.*

Theorem 2.2 *Under the Assumptions 2.1–2.6, 2.8 and the extra Assumption 2.9 there exists an optimal mechanism for the principal's utility maximization problem* (P)*. This same mechanism is also optimal for* (P^{as})*.*

In the proof, the optimal mechanism for (P) will come about by suitably modifying the optimal mechanism for (P^{as}) on the exceptional null set implicit in the definition of a.s. incentive compatibility. A useful sufficient condition for Assumption 2.8 to hold is as follows:

Proposition 2.1 *Suppose that the multifunction* Γ *has the following special form:*

$$\Gamma(t) := \{x \in K : V(t,x) \geq r(t)\},\ t \in T,$$

where $r : T \to \mathbf{R}$ *is some* reservation value *function (not necessarily measurable).*
a. Suppose that $\Gamma(t) \neq \emptyset$ *a.e. Then Assumption 2.7 follows from the other assumptions of Theorem 2.1, provided that Assumption 2.9 holds in addition.*
b. Suppose that $\Gamma(t) \neq \emptyset$ *for every* $t \in T$*. Then Assumption 2.8 follows from the other assumptions of Theorem 2.2.*

3 Applications

This section indicates how improvements of the main results of [9, 12, 13] all follow from Theorem 2.1 and how an improvement of the main result of [10] follows from Theorem 2.2. Since it would be most uneconomical to repeat here all the details of the models used in those papers, let us restrict ourselves to the main features of the implementation of the abstract existence results. Also, for each paper the improvements will be mentioned which follow from adopting Theorem 2.1 or 2.2. The reader is expected to fill out the details; in doing so, she/he is warned to observe the fact that no uniform terminology exists concerning the fundamental notions in this area. For instance, "incentive compatibility" in [9] is our present a.s. incentive compatibility (Definition 2.1), and "a.e. incentive compatibility" in [7] is a notion that is weaker than our present a.s. incentive compatibility.

In several applications the contracts will be *mixed*, i.e., K consists of the set $\mathcal{P}(Y)$ of all probability measures on some topological space Y, equipped with its Borel σ-algebra. In this situation E is taken to be the set of all bounded signed measures on Y, equipped with the classical narrow topology [5, III.54]. The following propositions are then useful for verifying the various assumptions used in Theorems 2.1 and 2.2.

Proposition 3.1 *Suppose that* Y *is a compact metric space and that* $t \mapsto Y(t)$ *is a multifunction from* T *into* 2^Y *with closed values. Then Assumptions 2.1, 2.2 hold for* $K := \mathcal{P}(Y)$ *and for* $\Gamma : T \to 2^K$ *given by*

$$\Gamma(t) := \mathcal{P}(Y(t)) := \{\pi \in \mathcal{P}(Y) : \pi(Y(t)) = 1\}.$$

Proposition 3.2 *Let* Y *and* $Y(t)$*,* $t \in T$*, be as above. Suppose* $v : T \times Y \to \mathbf{R}$ *and* $u : T \times Y \to [-\infty, +\infty)$ *are such that*

$$v(t,\cdot) \textit{ is bounded and continuous on } Y(t) \textit{ for each } t \in T,$$
$$u(t,\cdot) \textit{ is upper semicontinuous on } Y(t) \textit{ for each } t \in T,$$
$$u \textit{ is } \mathcal{T} \times \mathcal{B}(Y)\textit{-measurable},$$
$$u(t,y) \leq \psi(t) \textit{ for all } t \in T,$$

for some integrable function $\psi : T \to \mathbf{R}$. Then Assumptions 2.3–2.6 hold for $V : T \times \mathcal{P}(Y) \to \mathbf{R}$ and $U : T \times \mathcal{P}(Y) \to [-\infty, +\infty)$ given by

$$V(t,\pi) := \int_{Y(t)} v(t,y)\pi(dy),\ U(t,\pi) := \int_{Y(t)} u(t,y)\pi(dy).$$

Moreover, if also

$$v \text{ is } \mathcal{T} \times \mathcal{B}(Y)\text{-measurable,}$$

then Assumption 2.9 holds.

Proof of Proposition 3.1. All convexity considerations are trivial, so Assumption 2.1 follows by [5, III.60] and Assumption 2.2 follows by [5, III.58, III.60], since each $Y(t)$ is compact. QED

Proof of Proposition 3.2. As before, all convexity/affinity/concavity considerations are trivial. Therefore, validity of Assumption 2.3 follows directly from the definition of the narrow topology. Similarly, Assumption 2.6 holds directly by [5, III.55]. As for Assumption 2.4, consider first the case where u is the characteristic function of some set in $\mathcal{T} \times \mathcal{B}(Y)$. In this case the desired result follows directly from the proof of Theorem IV.12 in [4, pp. 103-104]. But once this is known, the usual approximation of u by means of step functions shows that the desired result is also valid for general u. Also, observe that Assumption 2.5 follows very simply from the given inequality for u Under the additional condition for v, validity of Assumption 2.9 is proven in almost the same way as Assumption 2.4, by applying the approximation by step functions to the positive and negative part of v. QED

3.1 Stackelberg games [9]

In [9] Page considers a Stackelberg game with incomplete information in the mixed contract setting of Propositions 3.1 and 3.2. Also, his definition of individual rationality is as given in Proposition 2.1. Improving upon Page's existence result for a problem of type (P^{as}), we shall show that his conditions also allow for an existence result as in Theorem 2.2. In the notation of [9], one should substitute for $(T, \mathcal{T}, \nu)$ its probability space (Θ, Σ, m), for K the set $P(X)$ of all probability measures on its compact metric space X, for $\Gamma(t)$ its sets $\mathcal{P}(X(t))$, where $X(\cdot)$ is its multifunction defined on its p. 415 (it is of the kind postulated in Proposition 2.1). Further, we should substitute for v its function f^*, defined on p. 414, and for u its function g^*, as given on p. 418 (correspondingly, our V and U are called F^* and G^* in [9]). We use simply $F(t) := T = \Theta$ in the definition of $\mathcal{S}_{IC}$. Let us check that the conditions of Theorem 2.2 apply. Above, we saw that the conditions of Proposition 3.1 hold; also, those of Proposition 3.2 apply: the needed properties of $v = f^*$ can be found on p. 414 (including the additional measurability) and those of $u = g^*$ on p. 419. By [A-2] in [9], Proposition 2.1.b applies. Therefore, the existence of an optimal mechanism for (P) follows. This improves upon [9, Theorem 5.1], where only the existence of an optimal solution for (P^{as}) is proven.

3.2 Principal-single agent models with incomplete information [10]

In [10] a principal-agent model is given with incomplete information about a single agent. Theorem 2.2 applies to this model, and this leads to at least one improvement. In [10], one should substitute $(T, \mathcal{T}, \nu)$ for its probability space $(\Theta, \mathcal{B}(\Theta), m)$. In [A-1] of [10], Θ is supposed to be a separable metric and complete space, but here we shall not need this topological assumption. We again use here $F(t) \equiv T$. Proposition 2.1.b applies here; its condition is fulfilled by [A-4] in [10] (note that it would already be enough to have the slightly lighter condition of Proposition 2.1.a hold here). Also, Proposition 3.1 applies to [10]; its conditions are seen to hold by p. 327 and p. 330 of [10], where one should substitute for our Y Page's space K and for $Y(t)$ Page's $K(t)$. Further, we may invoke Proposition 3.2, with the substitutions $v := V^*$ and $u := U^*$, in view of p. 330 and p. 332 of [10].

3.3 Principal-multiple agent models with incomplete information: mixed contracts [12]

In [12] a multi-agent contract auction model is considered. As announced in section 2, this type of model can be incorporated into the framework of section 2 by an adroit choice of the misspecification sets and the σ-algebra $\mathcal{T}$. Here the fact that Theorem 2.2 does not ask for measurability conditions for V, but only for U, plays a significant role. To capture Page's existence result, we substitute for T Page's $I \times \Theta$ (his I indexes the agents, and $\Theta := \Pi_{i \in I} \Theta_i$ is the set of all type-vectors). This T should be equipped with the special product σ-algebra $\mathcal{T} := \{\emptyset, I\} \times \mathcal{B}(\Theta)$. Observe that this has the effect of making $\mathcal{T}$-measurable functions depend essentially on the θ-part only of $t = (i, \theta)$. The measure ν is given by $\nu(I \times B) := \mu(B)$, where μ is the given measure on Θ as in [12]. Let us already note that Page uses topological conditions for Θ – of the kind specified in the previous subsection – that are not really needed in our approach. To capture the multi-agent setup entirely, we now define for $t = (i, \theta)$

$$F(t) := \{(i, \theta_{-i}, \theta'_i) : \theta'_i \in \Theta_i\}.$$

where $\theta^{-i} := (\theta_j)_{j \neq i}$. Thus, given i and the type-vector θ, only agent i is allowed to misreport her/his type. Otherwise, Page works again with the reservation value setup, described in Proposition 2.1. He also works with a mixed contract setup, which is now slightly at variance with the one described in Propositions 3.1–3.2. Namely, in his model we can substitute $Y := I \times \Phi$, $v := G^*$, $u := F^*$, and both his G^* and F^* are $\mathcal{T} \times \mathcal{B}(Y)$-measurable (they depend only upon the part θ in $t = (i, \theta)$). As for U, in [12] its definition is as in Proposition 3.2, and the conditions of that proposition are easily seen to hold (see p. 16 and Proposition 3.2.1 of [12]). But in the definition of V the individual's index i reappears:

$$V((i, \theta), \pi) := \int_{\{i\} \times \Phi} v(\theta, y) \pi(dy).$$

Of course, this destroys $\mathcal{T}$-measurability for V, but, as observed above, we do not need it for Theorem 2.1. But even though V is not exactly as in Proposition 3.2, continuity of V in its second variable, as desired in Assumption 2.3, can still easily be proven because of the discrete character of the topology on the first component of $Y := I \times \Phi$ (see p. 16 of [12]). Proposition 3.1 applies completely to [12], quite similar to the previous examples. Finally, Proposition 2.1, with its appeal to Assumption 2.9, fails to hold. This has been counteracted by stepping up the nonemptiness assumption in [12]: it is easy to see that [A-4] of [12], which asks for a *constant* function to lie in $\mathcal{S}_\Gamma$, causes that same constant also to belong to $\mathcal{S}_{IC}$.

3.4 Principal-multiple agent models with incomplete information: pure contracts [13]

In [13], a pure variant of the auction model in [12] was given. To capture and improve the main existence result in [13], we define $(T, \mathcal{T}, \nu)$ and the misspecification sets $F(t)$ as in Example . In the notation of [13], we now substitute $K := \Delta \times C$ and $\Gamma(t) := K(t)$, with $K(i, \theta) := \{(p, \xi) \in \Delta \times C : p_i v_i(\theta) - \xi_i \geq 0\}$. These satisfy Assumptions 2.1 and 2.2 by the compactness conditions on p. 4 of [13]. [1] Also, in the notation of [13], we should substitute $V(i, \theta, (p, \xi)) := p_i v_i(\theta) - \xi_i$. and $U(i, \theta, (p, \xi)) := u(v_0(\theta)(1 - \sum_j p_j)) + \sum_j \xi_j$. We observe that Assumption 2.3 is evident and that Assumption 2.9 does not hold, as happened in the previous subsection. Note how Page's assumption that u be concave and nondecreasing implies that $U(i, \theta, \cdot)$ is concave, as it should be in Assumption 2.6. In this case the structure of V is so simple that Assumption 2.7 holds instantly: any constant function identically equal to $(\bar{p}, 0) \in \Delta \times C$ lies in $\mathcal{S}_\Gamma \cap \mathcal{S}_{IC}$. Applying Theorem 2.1 gives a slight generalization of the main existence result of [13] for (P^{as}): once again, the topological conditions for Θ, imposed in [13], turn out to be redundant.

3.5 Principal-single agent models with incomplete information: pure state-contingent contracts [3]

In [3] Balder and Yannelis consider a model with state-contingent contracts, suggested by F.H. Page. In terms of their paper, we should make the following substitutions: $K := L^1_X$, $V := I_{U_2}$, $U := I_{U_1}$. They use a reservation value setup for Γ as in Proposition 2.1. Since they already refer to the present paper to deduce their main result, the reader can find in [3] a check-list of all the verifications of the assumptions used in Theorem 2.1.

4 Proofs of the main results

The proofs of Theorems 2.1, 2.2 are quite nontrivial and nonstandard because the delicate nature of the incentive compatibility constraint requires special attention. Those proofs are based on ideas of Page and involve two auxiliary results. The first of these is the following specialization of an abstract Komlós-type result, given in [1, Theorem 2.1].

Theorem 4.1 *i. Let (f_k) be a sequence of mechanisms in $\mathcal{S}_\Gamma^{as}$. Under Assumptions 2.1, 2.2 and 2.6 there exist a subsequence (f_m) of (f_k) and a mechanism f_* in $\mathcal{S}_\Gamma^{as}$ such that* [2]

$$\lim_{n \to \infty} \frac{1}{n} \sum_{m=1}^{n} f_m(t) = f_*(t) \ a.e.$$

ii. Moreover,

$$\limsup_k I_U(f_k) \leq I_U(f_*).$$

Proof. We shall apply Theorem 2.1 of [1]. In view of Assumption 2.1, it follows by [4, III.31] that there exists a countable subset (x_j^*) of E^* which separates the points of K.

[1] Surely C in [13] must be supposed convex; otherwise, it is impossible to see where the convexity conclusion for $D \cap \Gamma$ in [13, Theorem 2.2.3] should come from.

[2] This pointwise convergence of the averages continues to hold – with the same f_* but with varying exceptional null sets – for *every* subsequence of subsequence (f_m).

Define $a_j(t,x) := < x, x_j^* >$ on $T \times E$. For h as in [1, Theorem 2.1] we simply use $h := 0$ on $T \times K$ and $h := +\infty$ on $T \times (E \backslash K)$ (indeed, $\sup_k I_h(f_k) = 0 < +\infty$ and Assumption 2.1 causes $h(t,\cdot)$ to be trivially inf-compact on E for each t). Note also that Assumption 2.1 implies that $\gamma_j := \sup_{x \in K} < x, x_j^* >$ is finite for each j. Hence, property (B) of [1] holds. Thus, we may apply Theorem 2.1 of [1]. This gives the existence of a subsequence (f_m) and a function f_* such that the desired convergence statement holds. By [5, I.15], f^*, being the pointwise limit of a sequence of measurable functions, is measurable. By Assumption 2.2 it also follows easily that f^* belongs to $\mathcal{S}_\Gamma$. Part ii follows directly by [1, Corollary 2.2]. QED

Recall that a compact-valued multifunction $\Delta : T \to K$ is said to be *measurable* [4, III.2] if for every open or closed $O \subset K$ the set

$$\Delta^-(O) := \{t \in T : \Delta(t) \cap O \neq \emptyset\} \in \mathcal{T}.$$

Lemma 4.1 *Let (g_n) be a sequence of mechanisms $g_n : T \to K$. Under Assumption 2.1 the multifunction $L \to 2^K$ given by*

$$L(t) := \cap_{p=1}^\infty \textit{ closure } \{g_n(t) : n \geq p\},$$

is measurable.

Proof. Start by noting that for each $t \in T$ one has $x \in L(t)$ if and only if there exists a subsequence of $(g_n(t))$ converging to x. Let $O \subset K$ be an arbitrary open set. Then $t \in L^-(O)$ if and only if there exist $m \in \mathbf{N}$ and a subsequence of $(g_n(t))$ contained entirely in the set $O_m := \{x \in K : \text{dist}(x, K \backslash O) > 1/m\}$. [3] Indeed, if $t \in L^-(O)$ then there exists $x \in L(t) \cap O$. For some $m \in \mathbf{N}$ this gives $x \in L(t) \cap O_m$, so the desired subsequence is easily found by what was said in the first line of this proof, since O_m is open. Conversely, if there exist $m \in \mathbf{N}$ and a subsequence of $(g_n(t))$ contained in O_m, then, by Assumption 2.1, a further subsequence will converge to some x in the closure of O_m (i.e. $\text{dist}(x, K \backslash O) \geq 1/m$), and this closure is contained in O. We conclude from this characterization that

$$L^-(O) = \cup_m \cap_p \cup_{n \geq p} \{t \in T : f_n(t) \in O_m\},$$

which demonstrates that $L^-(O) \in \mathcal{T}$. QED

The above proof is from first principles; the result also follows by applying [4, III.4].

Proof of Theorem 2.1. By hypothesis the feasible set $\mathcal{S}_\Gamma^{as} \cap \mathcal{S}_{IC}^{as}$ for (P^{as}) is nonempty. So there exists certainly a maximizing sequence for (P^{as}). That is to say, there exists a sequence (f_k) in $\mathcal{S}_\Gamma^{as} \cap \mathcal{S}_{IC}^{as}$, with $\lim_k I_U(f_k) = \sup(P^{as})$. Application of Theorem 4.1 to (f_k) gives the existence of a null set N in $\mathcal{T}$, a subsequence (f_m) and $f_* \in \mathcal{S}_\Gamma^{as}$ such that $f_*(t) = \lim_n s_n(t)$ for every $t \notin N$. Here $s_n(t) := \frac{1}{n} \sum_{m=1}^n f_m(t)$. By part ii of Theorem 4.1,

$$I_U(f_*) \geq \sup(P^{as}). \tag{4.1}$$

Let us redefine $f_*(t)$ on N: consider the multifunction $L : N \to 2^K$, given by

$$L(t) := \cap_{p=1}^\infty \text{ closure } \{s_n(t) : n \geq p\}$$

[3] Here "dist" refers to distance with respect to a fixed metric on K.

By Lemma 4.1, this multifunction (which we need to consider only on N) is measurable. Evidently, by Assumption 2.1 the values of L are nonempty (use the finite intersection property) and compact. Therefore, L has a measurable selection $f_* : N \to K$ [4, III.6]. In view of (4.1), all that is needed to finish the proof now is to check that f_*, thus altered on N, is a.s. incentive compatible. Note that by incentive compatibility of the f_m's and by Assumption 2.3

$$V(t, s_n(t)) \geq V(t, s_n(t')) \text{ for all } t \in T \backslash N', \text{ all } t' \in F(t) \text{ and all } n \in \mathbf{N}. \tag{4.2}$$

Here N' is a null set; it is the union of all exceptional null sets involved in the a.s. incentive compatibility definition of the f_m's. Let $t \in T \backslash (N' \cup N)$ and $t' \in F(t)$ be arbitrary. If $t' \notin N$, then both $s_n(t) \to f_*(t)$ and $s_n(t') \to f_*(t')$. Thus, by (4.2) and Assumption 2.3

$$V(t, f_*(t)) \geq V(t, f_*(t')). \tag{4.3}$$

If $t' \in N$, then (4.3) holds as well. Indeed, by the alteration of f_* on N made above, $f_*(t')$ is a limit point of $(s_n(t'))$ (this is seen easily to be implied by the definition of the multifunction L; it was also observed in the proof of Lemma 4.1 above). Hence, there exists some subsequence (s_{n_j}) of (s_n) (its dependence on t' does not matter for the argument to go through) such that $(s_n(t'))$ converges to $f_*(t')$. But because $t \notin N$, the sequence $(s_{n_j}(t))$, obtained when the same subsequence is evaluated in the point t, converges to $f_*(t)$. Using (4.2) this gives

$$V(t, f_*(t)) \leftarrow V(t, s_{n_j}(t)) \geq V(t, s_{n_j}(t')) \to V(t, f_*(t)),$$

so (4.3) follows again. We therefore conclude that f_* is a.s. incentive compatible. QED

For the proof of Theorem 2.2 the following result is also needed. As shown here, when [4, III.39] is restricted to compact-valued multifunctions, it can be sharpened so as to give $\mathcal{T}$-measurability instead of measurability with respect to the ν-completion of $\mathcal{T}$. Although this is an obvious modification, we supply a proof that starts from first principles.

Lemma 4.2 *Let $\Delta : T \to 2^K$ be a measurable multifunction with nonempty compact values. Also, suppose that V satisfies Assumptions 2.3 and 2.9. Then the multifunction $\Delta_{\max} : T \to 2^K$, given by*

$$\Delta_{\max}(t) := \arg\max_{x \in \Delta(t)} V(t, x),$$

is measurable.

Proof. Of course, for every t

$$\Delta_{\max}(t) = \{y \in \Delta(t) : V(t, y) = \sup_{x \in \Delta(t)} V(t, x)\}.$$

Because Δ is measurable, it has a Castaing representation [4, III.9]. That is, there exists a countable subset $\{x_j : j \in \mathbf{N}\}$ of measurable selections of Δ such that

$$\Delta(t) = \text{closure of } \{x_j(t) : j \in \mathbf{N}\}$$

for each $t \in T$. Hence, by Assumption 2.3

$$V_{\max}(t) := \sup_{x \in \Delta(t)} V(t, x) = \sup_j V(t, x_j(t)) \text{ for all } t \in T.$$

Since each $t \mapsto V(t, x_j(t))$ is evidently $\mathcal{T}$-measurable, so is $V_{\max} : T \to \mathbf{R}$, defined above.

Now we can prove the desired measurability of $\Delta_{\max}$. Let $O \subset K$ be an arbitrary open set. For each $m \in \mathbf{N}$, let O_m be the set $\{x \in K : \operatorname{dist}(x, K \backslash O) > 1/m\}$; this set is open, and of course $\cup_m O_m = O$. So given $t \in \Delta^-(O)$, there must be $y \in \Delta(t)$ and $m \in \mathbf{N}$ such that $y \in \Delta_{\max}(t) \cap O_m$. It follows easily from the above Castaing representation of Δ and Assumption 2.3 that then for every $i \in \mathbf{N}$ there exists j such that

$$f_j(t) \in O_m \text{ and } V(t, f_j(t)) \geq V_{\max}(t) - \frac{1}{i}. \tag{4.4}$$

Conversely, if there exists n such that to every i there corresponds j with (4.4), then t must belong to $\Delta^-(O)$. Indeed, the corresponding subsequence of $f_j(t)$'s will have a convergent subsequence and a corresponding limit point z (by compactness of the set $\Delta(t)$). Assumption 2.3 then gives $V(t, z) = V_{\max}(t)$, since z lies in the set $\Delta(t)$. Hence, $z \in \Delta_{\max}(t)$. But the closure of O_m is still contained in O, so $z \in O$ as well. Together, these observations show that

$$\Delta^-(O) = \cup_n \cap_i \cup_j f_j^{-1}(O_m) \cap \{t \in T : V(t, f_j(t)) \geq V_{\max}(t) - \frac{1}{i}\}.$$

By the demonstrated measurability of $V_{\max}$ and the measurability of the f_j's this proves that $\Delta^-(O)$ belongs to $\mathcal{T}$. QED

Proof of Theorem 2.2. Analogous to the proof of Theorem 2.1 there exists a maximizing sequence (f_k) in $\mathcal{S}_\Gamma \cap \mathcal{S}_{IC}$, with $\lim_k I_U(f_k) = \sup(P)$. Again Theorem 4.1 gives the existence of a subsequence (f_m) and f_* in $\mathcal{S}_\Gamma^{as}$ such that for $s_n(t) := \frac{1}{n} \sum_{m=1}^n f_m(t)$ one has

$$s_n(t) \to f_*(t) \text{ for every } t \notin N. \tag{4.5}$$

Note already that by $(f_m(t)) \subset \Gamma(t)$ for all t we get $f_*(t) \in \Gamma(t)$ for $t \notin N$, by Assumption 2.2. Just as in the proof of Theorem 2.1, we also find that $I_U(f_*) \geq \sup(P)$. We shall alter f_* on N again, but in a different way than before (still following ideas of F. Page). Let $L : N \to 2^K$ be as defined in the proof of Theorem 2.1. This time, $L(t) \subset \Gamma(t)$ for *all* $t \in N$, by $(f_m(t)) \subset \Gamma(t)$ and Assumption 2.2. By Lemma 4.1, L is a measurable multifunction (on N). This fact allows us to invoke Lemma 4.2, and we find existence of a measurable function $\hat{f} : N \to K$ with $\hat{f}(t) \in \arg\max_{L(t)} V(t, \cdot) \subset L(t) \subset \Gamma(t)$ for all $t \in N$. Clearly, by defining $f_{**}(t) := f_*(t)$ for $t \in T \backslash N$ and $f_{**}(t) := \hat{f}(t)$ for $t \in N$, we obtain a function $f_{**} \in \mathcal{S}_\Gamma$ with $f_{**} = f_*$ a.e., whence $I_U(f_{**}) \geq \sup(P)$. It remains to check that f_{**} is incentive compatible. By Assumption 2.3, we know that

$$V(t, s_n(t)) \geq V(t, s_n(t')) \text{ for all } t \in T, \text{ all } t' \in F(t) \text{ and all } n \in \mathbf{N} \tag{4.6}$$

by incentive compatibility of the f_m's. Let $t \in T$ and $t' \in F(t)$ be arbitrary. Below we shall distinguish four different cases. If $t \notin N$, $t' \notin N$, then $s_n(t) \to f_*(t)$ and $s_n(t') \to f_*(t')$, so (4.6) implies

$$V(t, f_{**}(t)) \geq V(t, f_{**}(t')). \tag{4.7}$$

by Assumption 2.3. Secondly, if $t \notin N$ and $t' \in N$, then (4.7) holds as well. Indeed, $f_{**}(t') = \hat{f}(t')$ is still a limit point of $(s_n(t'))$, so there exists a subsequence $(s_{n_j}(t'))$ of $(s_n(t'))$ which converges to $f_{**}(t')$. But evaluating at $t \notin N$ also gives that $(s_{n_j}(t))$ converges to $f_*(t) = f_{**}(t)$ by (4.5). By (4.6)

$$V(t, f_*(t)) \leftarrow V(t, s_{n_j}(t)) \geq V(t, s_{n_j}(t')) \to V(t, f_{**}(t)),$$

so (4.7) follows. Conversely, if $t \in N$ and $t' \notin N$, then (4.7) holds as well. Indeed, it then reduces to $V(t, \hat{f}(t)) \geq V(t, f_*(t'))$, which follows by (4.6) and the "argmax" property of $\hat{t}$, since

$$V(t, f_{**}(t)) = V(t, \hat{f}(t)) \geq V(t, s_n(t)) \geq V(t, s_n(t')) \to V(t, f_*(t'))$$

again using Assumption 2.3. Fourthly, if both t and t' belong to N, then, working with the same subsequence $(s_{n_j}(t'))$ as above, we have by (4.6)

$$V(t, f_{**}(t) = V(t, \hat{f}(t)) \geq V(t, s_{n_j}(t)) \geq V(t, s_{n_j}(t')) \to V(t, f_{**}(t)).$$

So (4.7) has been shown to hold in all possible cases. This proves that f_{**} is incentive compatible. We conclude that f_{**} belongs to $\mathcal{S}_\Gamma \cap \mathcal{S}_{IC}$ and that $I_U(f_{**} = \sup(P)$. That f_{**} should also be optimal for (P^{as}) is an obvious consequence of its being the modification of f^*, the optimal solution for (P^{as}), on a null set. QED

Proof of Proposition 2.1. By Lemma 4.2 the multifunction $t \mapsto \arg\max_{x \in K} V(t, x)$ is measurable, and has nonempty values (the latter by the Weierstrass theorem). By the measurable selection theorem [4, III.6], this multifunction has a measurable selection $\tilde{f}$. We conclude that $V(t, \tilde{f}(t)) = \max_{x \in K} V(t, x)$ for all t. Evidently, this implies $\tilde{f} \in \mathcal{S}_{IC}$. Also, the nonemptiness hypothesis of part a gives $V(t, \tilde{f}(t)) \geq r(t)$ for a.e. t, and for part b the same inquality holds for all t. QED

5 Epilogue

It has been shown that the existence results for several models can be unified and derived from essentially two different (but related) existence results, viz., Theorems 2.1 and 2.2. To see the essential contours of Page's approach to the existence question for the class of models considered here should not only be useful in studying and classifying Page's work, but also to deal with new, more complex situations. Undoubtedly, the present model should be refined for that purpose. For instance, as yet the present author has not been able to deduce Page's recent existence results of [11], where a Bayesian multi-agent model is considered, from the present work.

References

[1] Balder, E.J. New sequential compactness results for spaces of scalarly integrable functions. *J. Math. Anal. Appl.* 151 (1990), 1-16.

[2] Balder, E.J. On equivalence of strong and weak convergence in L_1-spaces under extreme point conditions. *Israel J. Math.* 75 (1991), 21-47.

[3] Balder, E.J. and Yannelis, N.C. On necessary and sufficient conditions for continuity of expected utility functions. *Econ. Theory*, in press.

[4] Castaing, C. and Valadier, M. *Convex Analysis and Measurable Multifunctions.* Lecture Notes in Mathematics No. 580, Springer-Verlag, Berlin, 1977.

[5] Dellacherie, C. and Meyer, P.-A. *Probabilities and Potential.* North-Holland, 1978.

[6] Harsanyi, J.C. Games with incomplete information played by 'Bayesian' players. *Management Sci.* 14 (1967), 159-182.

[7] Khan, C.M. Existence and characterization of optimal employment contracts on a continuous state space. *J. Econ. Theory* 59 (1993), 122-144.

[8] Meyerson, R. *Game Theory.* Harvard University Press, 1991.

[9] Page, F.H. Incentive compatible strategies for general Stackelberg games with incomplete information. *Intern. J. Game Theory* 18 (1989), 409-421.

[10] Page, F.H. Optimal contract mechanisms for principal-agent problems with moral hazard and adverse selection. *Econ. Theory* 1 (1991), 323-338.

[11] Page, F.H. Bayesian incentive compatible mechanisms. *Econ. Theory* 2 (1992), 509-524.

[12] Page, F.H. Dominant strategy mechanisms for contract auctions with risk aversion and moral hazard. Preprint, June 1992.

[13] Page, F.H. Dominant strategy auction mechanisms. Preprint, July 1992.