Economic theory and methods

Taught by Eddie Dekel
Northwestern University, fall 2016

Ludvig Sinander
Northwestern University

This version: 28 March 2017

These notes are based on part of a theory course for second-year PhD students taught by Eddie Dekel at Northwestern in fall 2016. The topics are belief hierarchies and robustness.

I thank Eddie for teaching a great class and for agreeing to let me share these notes.

Contents

1   Hierarchies

Siniscalchi (2008) is a very good supplement to this section, especially on the interpretation of what we are doing. (Download the working paper version from Marciano's website, it's much better-typeset than the published version.)

1.1   Knowledge and common knowledge

The standard model of knowledge and belief in economics is as follows.[1] There is a set Ω of states of the world, equipped with a σ-algebra Σ. Each agent $i \in I$ is endowed with a map $\Pi_i : \Omega \to \Sigma$ such that $F_i(\Omega)$ is a partition of Ω. We call Π_i a partitional information function, or simply partition function.

   One topic that this framework allows us to address is that of common knowledge. For each $i \in I$, define the knowledge function $K_i : \Sigma \to \Sigma$ by

$$K_i(E) \coloneqq \{\omega \in \Omega : \Pi_i(\omega) \subseteq E\} \quad \forall E \in \Sigma.$$

(So $\omega \in K_i(E)$ reads 'i knows E at ω'.) For any event $E \in \Sigma$, define $E_0 \coloneqq E$, $E_n \coloneqq \bigcap_{i \in I} K_i(E_{n-1})$ for each $n \in \mathbf{N}$, and $C(E) \coloneqq \bigcap_{n \in \mathbf{N}} E_n$. We say that $E \in \Sigma$ is common knowledge at $\omega \in \Omega$ iff $\omega \in C(E)$, and we call $C : \Sigma \to \Sigma$ the common-knowledge function.

   Say that a set $F \in \Sigma$ is self-evident iff $\Pi_i(\omega) \subseteq F$ for every $\omega \in F$ and $i \in I$.[2] Aumann's (1976) original definition was that $E \in \Sigma$ is common knowledge at $\omega \in \Omega$ iff there is a self-evident set $F \in \Sigma$ such that $\omega \in F \subseteq E$. It is not hard to show that the two definitions are equivalent; see e.g. Osborne and Rubinstein (1994, Proposition 74.2).[3]

1.2   Why do we need a universal type space?

It seems sensible enough to read $K_j(E)$ and $K_i(F)$ as 'j knows that E' and 'i knows that F'. But $K_i(K_j(E))$ then reads 'i knows that $K_j(E)$'. If we want this to read 'i knows that j knows that E', then we have to assume that i recognises that $K_j(E)$ is equivalent to 'j knows that E'. But that requires that i 'knows' j's partition function Π_j.

   More generally, our definition of common knowledge can only be interpreted to mean that 'everyone knows that everyone knows that everyone knows...' if we assume that that partition functions $\{\Pi_i\}_{i \in I}$ are 'commonly

---

[1] See e.g. Osborne and Rubinstein (1994, ch. 5).

[2] Equivalently, F is self-evident iff $K_i(F) \supseteq F$ for every $i \in I$.

[3] Milgrom's (1981) axiomatic characterisation of common knowledge provides yet another equivalent definition.

known'. (Of course they cannot be common knowledge in the formal sense; hence the scare quotes.)

There is a view, usually attributed to Aumann (see e.g. his 1999), that it is without loss of generality to assume that partitions $\{\Pi_i\}_{i \in I}$ are 'common knowledge'. The argument is that if the partitions are not 'commonly known', then we need to consider a larger state space in which the states pin down what players know about each others' information.

It sounds like this might work, but (at least to me) it isn't obvious that it will. For example, the following approach will not work. If i and j have partitions $\in \Sigma^\Omega$ of Ω that are not 'common knowledge', expand the state space to $\Omega \times \Sigma^\Omega \times \Sigma^\Omega$. Unfortunately, there is no reason why the partitions of this enlarged state space should be 'common knowledge', either. We can continue expanding in this way as many times as we like, but there is no guarantee that we will eventually obtain 'common knowledge'.

The construction of a universal type space that we are about to carry out establishes that Aumann's intuition is in fact correct.[4] In particular, consider the enlarged state space $\Omega \times T^I$, where T is the universal type space. A state $\in \Omega \times T^I$ contains information both about $\omega \in \Omega$ and about the beliefs $\in T$ of each player. We will see that when it is 'common knowledge' that beliefs are coherent (not self-contradictory), we get 'common knowledge' of players' partitions of $\Omega \times T^I$ for free.

There is another motivation for what we are about to do. Consider a game of incomplete information, in which payoffs depend on a random element of some space S. In order to determine what she should do, player i must give consideration not just to her belief over S, but also to j's (joint) belief over i's belief over S and S, and so on. This directly leads us to consider an infinite hierarchy of beliefs.

Consider a seemingly different object: a particular kind of extensive-form game of imperfect information with a chance move, called a Bayesian game. In a Bayesian game, each player is endowed with a 'type space' (T_i, μ_i), where T_i is a set and $\mu_i(t_i)$ for each $t_i \in T_i$ is a belief over $S \times T_{-i}$. Nature draws the state and each player's type from $S \times \prod_{i \in I} T_i$. Each player learns (only) her own type $t_i \in T_i$, and forms a belief $\mu_i(t_i)$ over $S \times T_{-i}$ (possibly by updating a prior). The beliefs $\{\mu_i\}_{i \in I}$ clearly give rise to a belief hierarchy over S: beliefs over S, beliefs over $-i$'s types and hence $-i$'s beliefs over S, beliefs over $-i$'s beliefs over i's types and hence i's beliefs, and so on.

---

[4]There is another approach to deriving 'common knowledge' of partitions from more primitive assumptions. The idea is to define an epistemic logic, then to derive a state space and partitions from its sentences. This course is pursued by Aumann (1999); he calls it the 'syntactic' approach.

Harsányi (1967) argued informally that it is without loss to model games of incomplete information as Bayesian games. He reasoned that any belief hierarchy arising in a game of incomplete information can be generated by some type space in a Bayesian game; hence any incomplete-information game is strategically equivalent to some Bayesian game. This approach to studying games of incomplete information has been enormously fruitful!

But the claim that it is without loss of generality to study Bayesian games remains unproven. Of course we may choose the largest possible (i.e. universal) type space for each player, viz. the set T of all hierarchies. As with any type space, any belief over $S \times T^{|I|-1}$ generates a hierarchy $\in T$. But the formulation of a Bayesian game requires the inverse: an f that maps each hierarchy $\in T$ into a unique belief over $S \times T^{|I|-1}$. We will prove the existence (and uniqueness, and continuity, and more) of this inverse, and thereby put Harsányi's (1967) suggestion on firm footing.

1.3   Construction of the universal type space

The universal type space was first constructed by Armbruster and Böge (1979) and Böge and Eisele (1979);[5] both are mathematically difficult. The first construction that became well-known in economics was Mertens and Zamir (1985); this one is easier, but still uses some heavy machinery. We're going to follow Brandenburger and Dekel (1993), whose argument is much simpler and more transparent.

A topological space Z is Polish iff it is separable and completely metrisable. Note that a totally bounded Polish space is compact. For a Polish space Z, let $\Delta(Z)$ denote the set of measures on the Borel σ-field of Z, and endow it with the weak* topology (the topology of weak convergence). Then $\Delta(Z)$ is compact Polish.[6] We equip product spaces with the product topology and the associated Borel σ-field.

Let S be the basic space of uncertainty. Brandenburger and Dekel (1993) assume that S is Polish. For ease of exposition, further assume that S is totally bounded, hence compact. (We will discuss the topological assumptions later on.)

To avoid notational clutter, assume two players i and j. Define $X_0 = S$ and $X_n := X_{n-1} \times \Delta(X_{n-1})$ for $n \in \mathbf{N}$. We call $\delta_n \in \Delta(X_{n-1})$ an nth-order

---

[5]Bibliographical note: Armbruster and Böge (1979) is often erroneously cited as having been published in 1976. (Perhaps there was a working paper in 1976?)

[6]$\Delta(Z)$ is metrisable since the weak* topology is metrisable (by e.g. the Prohorov metric; see Billingsley (1999, pp. 72–3)). Separability and completeness are easy to show. Total boundedness is obvious, whence compactness.

belief. So $\delta_1 \in \Delta(S)$ is a first-order belief, $\delta_2 \in \Delta(S \times \Delta(S))$ is a second-order belief (joint belief about the state and j's belief about the state), and so on. Write $T_0 := \prod_{n=0}^{\infty} \Delta(X_n)$ for the set of all infinite hierarchies of beliefs $(\delta_1, \delta_2, \dots)$. T_0 is compact Polish.[7]

Consider a hierarchy $(\delta_1, \delta_2, \dots) \in T_0$. The nth-order belief for $n \geq 2$ is $\delta_n \in \Delta(X_{n-1}) = \Delta(X_{n-2} \times \Delta(X_{n-2}))$, and the $(n-1)$th-order belief is $\delta_{n-1} \in \Delta(X_{n-2})$. We define the hierarchy in this way in order to allow i's nth-order belief to have correlation between $x_{n-2} \in X_{n-2}$ and j's belief about x_{n-2}. But it involves redundancy: both δ_{n-1} and δ_n specify a marginal belief on X_{n-2} (and so does every higher-order belief). Rationality surely requires that these marginal distributions coincide.

Definition 1. $(\delta_1, \delta_2, \dots) \in T_0$ is coherent iff $\mathrm{marg}_{X_{n-2}} \delta_{n-1} = \delta_n$ for any $n \geq 2$.

We're looking for a universal type space $T \subseteq T_0$. As prefaced above, we'd like to choose T so that there is a map $f : T \to \Delta(S \times T)$ between types and beliefs over states and types that meets some criteria. The main criterion is that every type $(\delta_1, \delta_2, \dots) \in T$ maps into a belief $\in \Delta(S \times T)$ whose marginals equal δ_1, δ_2 and so on. We call this property 'belief-preservation':

Definition 2. A map $f : T_0 \to \Delta(S \times T_0)$ is belief-preserving iff for every $n \geq 1$, $\mathrm{marg}_{X_{n-1}} f((\delta_1, \delta_2, \dots)) = \delta_n$.

A second criterion is uniqueness: we'd like our f to be the only belief-preserving f.

Two further candidates are that f should be one-to-one and onto. But belief-preservation gives us both for free: if f is belief-preserving then its inverse is given explicitly by

$$f^{-1}(\delta) = \left(\mathrm{marg}_{X_0} \delta, \mathrm{marg}_{X_1} \delta, \mathrm{marg}_{X_2} \delta, \dots \right) \quad \forall \delta \in \Delta(S \times T_0).$$

Since we've given every set a topology, a further (but arguably optional) criterion is for f and its inverse to be continuous, so that it is a homeomorphism. As discussed below, it isn't clear why topology should matter, at least for our intuitive understanding of the universal type space. But conditional on having topological structure, we may agree that continuity is a good thing.

Define $T_1 := \{t \in T_0 : t \text{ is coherent}\}$. We're going to construct a belief-preserving map $f : T_1 \to \Delta(S \times T_0)$ that will be unique and a homeomorphism. The first (and main) step is the following.

---

[7]Metrisability, separability and compactness are all obvious.

Proposition 1. There is a unique belief-preserving $f : T_1 \to \Delta(S \times T_0)$, and it is a homeomorphism.

f has the properties we want: it is belief-preserving, unique and a homeomorphism. But we're not done: f tells us that types of i in T_1 can be associated uniquely with beliefs of i about S and j's type in T_0. The type spaces are not the same: i does not necessarily rule out that j's type is incoherent. Recall, however, that by a universal type space we mean $T \subseteq T_0$ such that there is an appropriate map from T to $\Delta(S \times T)$. Since T is on both the right- and left-hand sides, the model is 'closed': player i's type in T pins down a belief over S and player j's in the *same* T, and vice versa.

It turns out to be easy to get this 'closedness' property. Define

$$T_k := \{t \in T_1 : f(t)(S \times T_{k-1}) = 1\} \quad \text{for each } k \geq 2.$$

This imposes higher orders of belief in coherency, e.g. a type in T_2 is coherent and believes that her opponent is coherent. $\{T_k\}$ is a decreasing sequence of compact sets, so $T := \bigcap_{k \in \mathbf{N}} T_k$ is nonempty and compact. Intuitively, T is the set of types consistent with 'common knowledge' of coherency.

Consider the restriction of f to T. It is obviously still belief-preserving, unique, one-to-one, and continuous with a continuous inverse. Some reflection shows that it is also onto $\Delta(S \times T)$. For if $\delta \in \Delta(S \times T)$, then δ places zero probability on hierarchies that are not consistent with 'common knowledge' of coherency. The hierarchy $f^{-1}(\delta)$ that type δ generates must therefore be consistent with 'common knowledge' of coherency, i.e. $f^{-1}(\delta) \in T$ as desired. We have shown:

Proposition 2. There is a unique belief-preserving $f : T \to \Delta(S \times T)$, and it is a homeomorphism.

So T, the set of hierarchies consistent with 'common knowledge' of coherency, is the universal type space that we were after.

Let's fill in some of the holes in the argument. A fuller argument can be found in Brandenburger and Dekel (1993).

Sketch proof of Proposition 1. Consider $(\delta_1, \delta_2, \dots) \in T_1$, define

$$C_n := \left\{ \delta \in \Delta(S \times T_0) : \max_{X_{n-1}} \delta = \delta_n \right\},$$

and let $C := \bigcap_{n \in \mathbf{N}} C_n$ be the set of beliefs on $S \times T_0$ whose marginals coincide with $(\delta_1, \delta_2, \dots)$. C contains the measures whose value on the cyclinder sets

$\{C_n\}_{n \in \mathbf{N}}$ are pinned down by $(\delta_1, \delta_2, \dots)$. $\{C_n\}_{n \in \mathbf{N}}$ is a decreasing sequence of nonempty compact sets, so C is nonempty.[8] A standard result in measure theory says that C contains at most one element, roughly because the cylinders are a sufficiently rich class of sets that two measures cannot agree on the cylinders and yet disagree elsewhere.[9] So let $f((\delta_1, \delta_2, \dots))$ equal the unique element of C.

The $f : T \to \Delta(S \times T)$ thus constructed is obviously uniquely belief-preserving. As argued above, belief-preservation implies invertibility. To see that f^{-1} is continuous, simply observe that

$$f^{-1}(\delta) = \left(\underset{X_0}{\operatorname{marg}}\, \delta, \underset{X_1}{\operatorname{marg}}\, \delta, \underset{X_2}{\operatorname{marg}}\, \delta, \dots \right) \quad \forall \delta \in \Delta(S \times T_0)$$

and that marg is a continuous operator.[10] Continuity of f can be shown using a simple sequential argument that we omit. ∎

1.4 Varying the topological assumptions

The construction we just performed involved some assumptions. In particular, we assumed that S has topological structure, and further gave the weak* topology to sets of measures and the product topology to product spaces.

The weak* topology is arguably a natural choice because it implies that expected utilities are continuous in beliefs.[11] In some settings, though, it may be important to use a topology that makes the best-reply correspondence continuous. The weak* topology does not have this property.

Thse use of the product topology may also be considered natural. When T has the product topology, a sequence of belief hierarchies converges iff every level of the hierarchy converges separately, which seems like an intuitive notion of closeness.[12] But the product topology on T is not always a good choice. Notably, the rationalizability correspondence (for a fixed game) need not be continuous in the product topology.[13] Some care is needed!

---

[8]If we drop compactness, then we instead use Kolmogorov's existence theorem (for stochastic processes) to show that C is nonempty.

[9]The result that a measure is uniquely determined by its value on the cylinders can be viewed as a corollary of Dynkin's π-λ theorem (a much more general result). See e.g. Billingsley (1995, sec. 3).

[10]Proving that marg is continuous is trivial.

[11]This is obvious from the definition of weak convergence.

[12]Though the uniform topology can be motivated in the same way.

[13]Dekel, Fudenberg and Morris (2006) provide a 'strategic' topology on T in which the rationalizability correspondence is continuous. Their topology is finer than the product topology.

It is possible to construct universal type spaces using different topological assumptions. The best-known constructions are those of Mertens and Zamir (1985), who assume that S is compact, and Brandenburger and Dekel (1993), who assume that S is Polish (but not necessarily totally bounded). Heifetz (1993) constructs a universal type space assuming only that S is Hausdorff and that beliefs are regular. Mertens, Sorin and Zamir (2015, ch. III) provide results for a range of topological assumptions.

A bigger question is why we used topological structure at all. Intuitively, it seems that the universal type space for a given underlying space of uncertainty S should be a purely measure-theoretic construct: it's about beliefs (measures) and nothing else. But the construction above relied heavily on Kolmogorov's existence theorem, which uses topological structure. In fact, Kolmogorov's existence theorem is false for arbitrary measurable spaces! This raises the question of whether it is possible (using a different method of proof) to construct a universal type space without resort to topology.

The type space T we constructed above (i.e. the set of all belief hierarchies consistent with 'common knowledge' of coherency) is not a universal type space when no topological structure is imposed (Heifetz & Samet, 1999). However, Heifetz and Samet (1998) show that one can construct a (different) universal type space without making use of topological structure. (They make measurability assumptions, but no topological assumptions.) Meier (2012) proved that the Heifetz and Samet (1998) universal type space is complete.[14]

1.5   Extensions

Suppose we are interested in games in extensive form. In that case, we need to think of hierarchies of *conditional* beliefs.[15] In particular, each level of the hierarchy will be a regular conditional probability system rather than a single probability measure.[16] Battigalli and Siniscalchi (1999) construct a universal type space of regular conditional probability systems (for a given σ-algebra of conditioning events).

---

[14]Completeness was a by-product of how we constructed our universal type space T above, but that is not the case for the Heifetz and Samet (1998) construction.

[15]If everything is on-path then the conditional probabilities are pinned down by the prior probabilities, so we can use the previous construction. But extensive-form games generally require us to consider probabilities following events of prior measure zero.

[16]The notion of a regular conditional probability is due to Rényi (1955), and was brought to economics by Myerson (1986). See Myerson (1991, sec. 1.6) for a (brief) textbook treatment.

Another class of extensions allows for more general notions of 'belief'. Ahn (2007) constructs a complete and universal type space for compact multiple-prior models: $f : T \to K(\Delta(S \times T))$, where $K(Z)$ denotes the set of all compact subsets of Z.

Epstein and Wang (1996) construct a type space in an even more general setting that does not necessarily even involve probabilities. In particular, they construct a T and a well-behaved map $f : T \to P(S \times T)$, where $P(S \times T)$ is the set of all acts. Some restrictions on preferences are of course needed for this to work.[17] The authors do not show that their T is either universal or complete. Chen (2010) shows that it is in fact universal and complete (under suitable assumptions), and that some of Epstein and Wang's technical conditions can be dropped.

All of these papers construct spaces of belief hierarchies. Mariotti, Meier and Piccione (2005) show how to construct a space of knowledge hierarchies: in particular, a type space T and a suitable map $f : T \to K(S \times T)$. They show that their T is universal and complete.

1.6 Applications

Epstein and Peters (1999) provide a canonical class of mechanisms (i.e. a revelation principle) for mechanism design environments with multiple principals.[18] Suppose we have principals i and j offering direct revelation mechanisms to an agent whose type is privately known. Suppose i's mechanism asks the agent to report her type and her report to j. If j also offers such a mechanism, then i may be able to gain by using a 'second-order' mechanism that asks for the state, her report to j about the state, and her report to j about her report to i. To construct a universal space of mechanisms, the authors proceed as we did above, making use of a coherency condition similar to ours. This is interesting, but not obviously useful. Their revelation principle says that we can restrict attention to a certain class of mechanisms, but that class is enormous!

Lipman (1991) studies decision problems. To decide, the agent first has to decide how to decide, but that in turn requires her to decide how to decide how to decide, and so on. Under some assumptions, it is possible to construct a universal space of decision problems. Like the last paper, this exercise is illuminating, but probably not useful.

---

[17]In particular, while Ahn's (2007) result is a special case 'in spirit', his multiple-priors setting is technically ruled out by Epstein and Wang's (1996) assumptions.

[18]The multiple-principals setting is sometimes called 'common agency'.

There are more useful applications, however! One notable such is Epstein and Zin (1989), who extend Kreps and Porteus's (1978) work on preferences for the timing of the resolution of uncertainty to an infinite-horizon context. This extension is not so easy, as it turns out. First, an agent has preferences over consumption today and lotteries over consumption streams that resolve tomorrow: $D_1 = \mathbf{R} \times \Delta(\mathbf{R}^\infty)$.[19] Second, she has preferences over consumption today and lotteries that resolve tomorrow over lotteries that resolve the following day: $D_2 = \mathbf{R} \times \Delta(D_1)$. Continuing in this fashion, we have a 'tth-order' choice set $D_t = \mathbf{R} \times \Delta(D_{t-1})$ in which all uncertainty is resolved in t periods' time. Preferences must be defined over a 'universal' choice set that includes all of these possible timings of the resolution of uncertainty. Epstein and Zin (1989) prove the existence of a universal and complete space D and a well-behaved map $f : D \to \mathbf{R} \times \Delta(D)$. Once D has been constructed, the authors extend to preferences defined on D Kreps and Porteus's (1978) axiomatisation of attitudes toward the timing of the resolution of uncertainty, and go on to study the asset-pricing implications of such preferences.

Another 'useful' application is Gül and Pesendorfer's (2016) infinite-horizon extension of their two-period temptation model (Gül & Pesendorfer, 2001). In this problem, preferences are defined over an 'infinitely regressing' sequence of spaces for roughly the same reason as in Epstein and Zin (1989).[20] This again requires the construction of a 'universal' space on which preferences can be defined. More generally, any infinite-horizon decision problem in which agents care about things like timing will require preferences to be defined on a space constructed along the lines of the universal type space.

Yet another useful application is Gül and Pesendorfer's (2016) model of interdependent preferences. They define an interdependent-preferences model as a triple (T, γ, ω), where T is the set of types, γ maps T^2 into preference profiles, and ω maps T into relevant physical characteristics (e.g. valuations).[21] The fact that γ is defined on T^2 means that each player's preferences depend on the other player's type. For example, a generous type of a player may be more willing to give if she encounters another generous person. This generates an infinite hierarchy of preferences, so the authors ask the natural question of whether every such infinite hierarchy can be mapped back into a 'universal' interdependent-preferences model. The answer turns out to be 'no': some interdependent-preferences models cannot

---

[19]I wrote $\mathbf{R}^\infty$ for the space of sequences, but the authors restrict attention to ℓ^p spaces.

[20]It isn't exactly the same reason, though: there is no uncertainty in (2016).

[21]The fact that ω is defined on T rather than T^2 means that the physical characteristics are not interdependent, thus ruling out settings such as common-values auctions that are not in fact instances of preference interdependence.

be recovered from the preference hierarchies they generate. Interestingly (and controversially), the argue that we should consider only the subset of models for which a universal model exists.

	A	B
A	M, M	$1, -L$
B	$-L, 1$	$0, 0$

(a) g_a (probability $1 - q$)

	A	B
A	$0, 0$	$1, -L$
B	$-L, 1$	M, M

(b) g_b (probability q)

Figure 1 – Payoffs in the email game. Assume $L > M > 1$ and $q < 1/2$.

2 Robustness

Game-theoretic analysis relies heavily on common knowledge. We might hope that conclusions continue to hold approximately when these common-knowledge assumptions are relaxed slightly. The robustness literature explores this question.

2.1 The email game

Rubinstein (1989) showed that robustness is not for free. His email game provides an example of a game of complete information and a class of incomplete-information perturbations such that a strict Nash equilibrium of the former is very different from rationalisable behaviour in any of the perturbed games.[22]

The email game is a Bayesian game with players 1 and 2, both with action set $\{A, B\}$. Payoffs are given by one of the two complete-information games g_a and g_b depicted in Figure 1, where the constants M and L satisfy $L > M > 1$. The common prior is that g_b is played with probability $q < 1/2$.

The unique Nash equilibrium of g_a is (A, A). (Indeed A is strictly dominant for each player.) g_b has two strict pure-strategy Nash equilibria: the 'safe' (A, A) and the 'risky' but Pareto-dominant (B, B). Preview: even when it is 'almost' common knowledge that payoffs are given by g_b, no type of either player will ever play the risky action B, so (B, B) is a not 'robust', even though it is a strict Nash equilibrium.

The information structure is as follows. Player 1 learns the payoffs, and player 2 does not. If payoffs are g_a then nothing happens. If payoffs are g_b then player 1 (automatically) sends an email to player 2, which arrives with (high) probability $1 - \varepsilon$. If player 2 gets an email from 1, she (automatically) sends an email back to confirm receipt, and this email arrives with probability $1 - \varepsilon$. This continues until an email is lost: whenever a player receives an

---

[22]The exposition here borrows from Osborne and Rubinstein (1994, sec. 5.5).

	n_2					
	0	1	2	3	4	$\cdots$
0	$1-q$					
1	$q\varepsilon$	$q\varepsilon(1-\varepsilon)$				
n_1 2		$q\varepsilon(1-\varepsilon)^2$	$q\varepsilon(1-\varepsilon)^3$			
3			$q\varepsilon(1-\varepsilon)^4$	$q\varepsilon(1-\varepsilon)^5$		
4				$q\varepsilon(1-\varepsilon)^6$	$q\varepsilon(1-\varepsilon)^7$	
$\vdots$				$\ddots$	$\ddots$	

Figure 2 – The common prior μ on the state space Ω of the email game.

email, she sends a response. (Except for 1's initial email, no-one ever sends an email except as a response.)

To formalise this, let n_i be the number of emails that player i has sent. We can use the state space

$$\Omega = \left\{ (n_1, n_2) \in \mathbf{Z}_+^2 : n_1 = n_2 \text{ or } n_1 = n_2 + 1 \right\}.$$

In state (n, n), player 2 received all n of player 1's messages, but her nth response got lost. In state $(n, n+1)$, player 1 received all n of player 2's messages, but her $(n+1)$th message got lost. Payoffs are g_a iff the state is $(0,0)$. The states and their probabilities are as depicted in Figure 2. Formally, the common prior μ is given by $\mu(0,0) = 1 - q$,

$$\mu(n+1, n) = q\varepsilon(1-\varepsilon)^{2n} \quad \text{and} \quad \mu(n+1, n+1) = q\varepsilon(1-\varepsilon)^{2n+1}$$

for each $n \in \mathbf{Z}_+$.

Players know how many messages they sent, but not whether the other player sent the same number of messages. In the graphical depiction in Figure 2, player 1 knows the row (n_1), and player 2 knows the column (n_2). Formally, the information functions are given by $\Pi_1(0,0) = \{0,0\}$,

$$\Pi_1(n+1, n) = \Pi_1(n+1, n+1) = \{\{n+1, n\}, \{n+1, n+1\}\},$$
$$\text{and} \quad \Pi_2(n+1, n) = \Pi_2(n, n) = \{\{n+1, n\}, \{n, n\}\}$$

for every $n \in \mathbf{Z}_+$.

A simple argument shows that (for $\varepsilon > 0$ sufficiently small,) there is a unique interim rationalisable strategy profile in which every type of each

player chooses A. Type $n_1 = 0$ of player 1 knows that payoffs are g_a, so must play A. Type $n_2 = 0$ of player 2 therefore knows that player 1 chooses A with probability at least $1 - q > 1/2$, so must play A. Type $n_1 = 1$ then knows that player 2 chooses A with probability at least $1/(2 + \varepsilon)$, so must play A. So type $n_2 = 1$ knows that player 1 chooses A with probability at least $1/(2 + \varepsilon)$, so plays A. Continuing by induction, we find that rationalisability requires every type of each player to choose A.

An iterated deletion argument with this structure is often called an 'contagion' (or 'infection') argument. The idea originates with Rubinstein (1989), and it is at the heart of the robustness and global games literatures. The key ingredient is the existence of 'dominance regions' such that some action is never a best reply to beliefs in such a dominance region. General negative robustness results often rely on showing the existence of dominance regions, then applying a contagion argument. A common way to obtain dominance regions is to assume 'payoff richness', meaning that the class of perturbations is sufficiently large that it contains games in which some actions are dominated for some types.

In the complete-information game g_b, payoffs are common knowledge, and (B, B) is a strict Nash equilibrium. In the email game, when payoffs are g_b, payoffs are 'almost' common knowledge, and yet B is never played. The sense in which payoffs are almost common knowledge is straightforward: conditional on payoffs being given by g_b, both players know that payoffs are g_b with high probability $1 - \varepsilon$, players know that the other player knows that payoffs are g_b with high probability $1 - \varepsilon \left[1 + (1 - \varepsilon) + (1 - \varepsilon)^2 \right]$, and so on.

A more formal sense in which payoffs are almost common knowledge is that as $\varepsilon \longrightarrow 0$, the players' types converge in the the product topology on the universal type space to a common-knowledge-of-payoffs pair of types. Convergence in the product topology simply means that for each $k \in \mathbf{N}$, kth order beliefs converge (separately). So we are taking 'almost common knowledge' to mean 'close in the product topology to common knowledge'. The fact that there is a sequence of perturbations along which B may not be played, but which converges to a game in which B may be played, is precisely a failure of hemicontinuity in the product topology of the correspondence from types to behaviour.

Rubinstein (1989) considers the non-robustness of a strict Nash equilibrium in the email game a paradox, or at least a problem. He thinks that (B, B) is a perfectly sensible prediction for g_b even if we imagine that players are slightly unsure about payoffs. (The story is, roughly speaking, that real people don't use mathematical induction when reasoning strategically.) The (cheap) way to get around this is to simply rule out these kinds of per-

turbations. A more principled reponse is to argue that precisely because of email-game-type pathologies, the product topology provides the wrong notion of closeness of types. Dekel et al. (2006) develop a topology on the universal type space under which the correspondence from types to rationalisable behaviour is lower hemicontinuous.[23]

The subsequent literature has instead tended to view the non-robustness of certain strict Nash equilibria to incomplete-information perturbations as good news. The refinements literature of the 1980s may allow us to refine away some Nash equilibria, but it is mostly silent on the question of how to select between multiple strict Nash equilibria.[24] By contrast, the email game shows that robustness to incomplete information is a refinement that may eliminate strict Nash equilibria. The global games literature, starting with Carlsson and van Damme (1993), is concerned with refinements of this sort.[25]

2.2 Common p-belief

The robustness literature can be viewed as an extended response to the email game. The goal is to understand when and why our predictions in complete-information games break down when common knowledge is relaxed slightly.

The email game appeared to have almost common knowledge of payoffs. But it turned out that the information structure was not 'close enough' to common knowledge to guarantee behaviour close to behaviour under common knowledge. This raises the question of whether there is a weakening of common knowledge that rules out such failures of lower hemicontinuity. Monderer and Samet (1989) answer this question in the affirmative: their notion of common p-belief has this property.

2.2.1 p-belief

Consider again the model of knowledge from section 1.1 (p. 4): a set of states Ω equipped with a σ-algebra Σ, and maps $\Pi_i : \Omega \to \Sigma$ for each player i such that $\Pi_i(\Omega)$ partitions Ω. The knowledge operator $K_i : \Sigma \to \Sigma$ is defined

$$K_i(E) := \{\omega \in \Omega : \Pi_i(\omega) \subseteq E\} \quad \forall E \in \Sigma.$$

---

[23]In particular, their 'strategic topology' is the coarsest topology with this property.

[24]For example, the Kohlberg–Mertens (1986) stable set contains every strict Nash equilibrium.

[25]The strange term 'global game' comes from Carlsson and van Damme (1993).

Now add to the model a belief for each agent, i.e. a probability measure μ_i on (Ω, Σ) for each $i \in I$. Assume that each agent i's prior μ_i assigns positive measure to each cell of every agent j's partition $\Pi_j(\Omega)$ of Ω, so that conditional probabilities are unique.[26] (Note that this implies that the partitions are countable.) For $p \in [0, 1]$, define $B_i^p : \Sigma \to \Sigma$ by

$$B_i^p(E) := \{\omega \in \Omega : \mu_i\left(E | \Pi_i(\omega)\right) \geq p\} \quad \forall E \in \Sigma$$

$B_i^p(E)$ is the event that i believes E with probability at least p; we say that 'i p-believes E.

$B_i^1(E)$ and $K_i(E)$ do not coincide when there are null sets. In particular, while K_i satisfies $K_i(E) \subseteq E$ (the 'axiom of truth'), $B_i^1(E)$ does not imply that E obtains. (Think of Lebesgue measure on $[0, 1]$.) But they do coincide whenever Ω is countable. This is the only important thing that changes when Ω is allowed to be uncountable; we will therefore assume Ω countable (and that $\Sigma = 2^\Omega$) to make our lives easier.

B_i^p has the following easily-derived properties.

Proposition 3. For any $p, q \in [0, 1]$ and $E, F \in \Sigma$,

(1) $B_i^p\left(B_i^p\left(E\right)\right) = B_i^p(E)$.[27]

(2) If $p > q$ then $B_i^p(E) \subseteq B_i^q(E)$.

(3) If $E \subseteq F$ then $B_i^p(E) \subseteq B_i^p(F)$.

(4) $B_i^p(E \cap F) \subseteq B_i^p(E) \cap B_i^p(F)$, with equality if $E \subseteq F$.[28]

(5) $B_i^p(E)$ is a union of the cells of the partition $\Pi_i(\Omega)$.

With the exception of (2), K_i satisfies all of these properties. But K_i satisfies property (4) with equality even when E and F are not nested, whereas B_i^p does not. To see why, suppose that E and F are disjoint, that each occurs with probability $1/2$, and that $p = 1/3$. Then $B_i^p(E \cap F) = \varnothing$, but $B_i^p(E) \cap B_i^p(F) = \Omega \cap \Omega = \Omega$.

---

[26] The results below (most of which are from Monderer and Samet (1989)) can be extended to the case in which some partition cells have measure zero, but it is somewhat delicate since conditional probabilities are no longer unique. See Kajii and Morris (1997a).

[27] This can be strengthened to 'if $q \geq p$ then $B_i^q\left(B_i^p\left(E\right)\right) = B_i^p(E)$'.

[28] This can be strengthened to $B_i^p\left(\bigcap_{n \in \mathbf{N}} E_n\right) = \bigcap_{n \in \mathbf{N}} B_i^p(E_n)$ if $E_n \subseteq E_{n-1} \ \forall n \in \mathbf{N}$.

2.2.2 Common p-belief

When we talk about multiple agents, it becomes important that they do not assign probability zero to different events. So assume that they do not, i.e. that $\{\mu_i\}$ are mutually absolutely continuous.[29] It is then wlog to assume that the priors have full support on Ω (just throw away any zero-probability states), so let's do that. To lighten the notational burden, assume two players i and j.

Recall how we defined common knowledge in section 1.1 (p. 4). Our primary definition was iterative: for any event $E \in \Sigma$, define $E_0 := E$, $E_n := K_i(E_{n-1}) \cap K_j(E_{n-1})$ for each $n \in \mathbf{N}$, and $C(E) := \bigcap_{n \in \mathbf{N}} E_n$; we say that $E \in \Sigma$ is common knowledge at $\omega \in \Omega$ iff $\omega \in C(E)$. We also gave an equivalent definition due to Aumann: $E \in \Sigma$ is common knowledge at $\omega \in \Omega$ iff there is a self-evident set $F \in \Sigma$ such that $\omega \in F \subseteq E$, where a set $F \in \Sigma$ is self-evident iff $K_i(F) \supseteq F$ for every and $i \in I$.

Common p-belief is defined analogusly to common knowledge: for any event $E \in \Sigma$, define $E_0 := E$,

$$E_n := B_i^p(E_{n-1}) \cap B_j^p(E_{n-1}) \quad \text{for each } n \in \mathbf{N},$$

and $C^p(E) := \bigcap_{n \in \mathbf{N}} E_n$, and say that $E \in \Sigma$ is common p-belief at $\omega \in \Omega$ iff $\omega \in C^p(E)$.

A nice feature of this concept is that it has a nice Aumann-style 'fixed-point' characterisation. Call a set $F \in \Sigma$ p-evident at $\omega \in \Omega$ iff $B_i^p(F) \supseteq F$ for each player $i \in I$.

Proposition 4. $\omega \in C^p(E)$ iff there is a p-evident set $F \in \Sigma$ such that $\omega \in F \subseteq B_i^p(E)$.

The proof is straightforward: see Monderer and Samet (1989, p. 177–8).

Morris (1999) considers various alternative 'iterative' (he calls them 'hierarchical') definitions of approximate common knowledge; we'll look at one of them. Define $I^p : \Sigma \to \Sigma$ by

$$I^p(E) := \left[B_i^p(E) \cap B_i^p\left(B_j^p(E)\right) \cap B_i^p\left(B_j^p(B_i^p(E))\right) \cap \cdots \right]$$
$$\cap \left[B_j^p(E) \cap B_j^p(B_i^p(E)) \cap B_j^p\left(B_i^p\left(B_j^p(E)\right)\right) \cap \cdots \right]$$

---

[29]This imposes some degree of agreement among agents. We might argue that it is a weak restriction when Ω is countable, since each singleton can then be assigned (possibly very small) positive probability by every player. But for Ω uncountable, there's no getting around the fact that mutual absolute continuity of priors is a strong assumption. (Though not as strong as assuming a common prior, obviously.)

for every $E \in \Sigma$. Say that $E \in \Sigma$ is iterated p-belief at $\omega \in \Omega$ iff $\omega \in I^p(E)$. (The difference is in the order of intersection.) For $p = 1$, common and iterated p-belief coincide.[30] For $p < 1$, it's easy to show that $C^p \subseteq I^p$ (see Morris (1999, Lemma 14)). But the containment may be strict in general, as the following example from Morris (1999) demonstrates.

Example 1. Consider $\Omega = \{1, 2, 3, 4, 5, 6\}$, the partitions

$$\Pi_i(\Omega) = \{\{1, 2\}, \{3\}, \{4\}, \{5, 6\}\}$$
$$\Pi_j(\Omega) = \{\{1, 3, 4\}, \{2, 5, 6\}\},$$

and the uniform common prior $\mu_i(F) = \mu_j(F) = |F|/6$ for each $F \subseteq \Omega$. Let $p = 0.6$ and $E = \{1, 2, 3\}$. We're going to show that $C^p = \varnothing$ and $I^p = \{3\}$, so that the containment $C^p \subseteq I^p$ is strict in this case.

$B_i^p(E) = \{1, 2, 3\}$ and $B_j^p(E) = \{1, 3, 4\}$, so

$$E_1 = B_i^p(E) \cap B_j^p(E) = \{1, 3\}.$$

Then $B_i^p(E_1) = \{3\}$ and $B_j^p(E_1) = \{1, 3, 4\}$, so

$$E_2 = B_i^p(E_1) \cap B_j^p(E_1) = \{3\}.$$

$B_i^p(E_2) = \{3\}$ and $B_j^p(E_2) = \varnothing$, so

$$E_3 = B_i^p(E_2) \cap B_j^p(E_2) = \varnothing.$$

Hence $C^p(E) = \bigcap_{n \in \mathbf{N}} E_n = \varnothing$.

Again $B_i^p(E) = \{1, 2, 3\}$ and $B_j^p(E) = \{1, 3, 4\}$. Hence

$$B_i^p\big(B_j^p(E)\big) = B_i^p(\{1, 3, 4\}) = \{3, 4\}$$
$$\text{and} \quad B_j^p\big(B_i^p(E)\big) = B_j^p(\{1, 2, 3\}) = \{1, 3, 4\}.$$

So

$$B_i^p\big(B_j^p(B_i^p(E))\big) = B_i^p(\{1, 3, 4\}) = \{3, 4\}$$
$$\text{and} \quad B_j^p\big(B_i^p(B_j^p(E))\big) = B_j^p(\{3, 4\}) = \{1, 3, 4\}.$$

Spot the pattern? We therefore have

$$B_i^p(E) \cap B_i^p\left(B_j^p(E)\right) \cap B_i^p\left(B_j^p(B_i^p(E))\right) \cap \cdots$$
$$= \{1, 2, 3\} \cap \{3, 4\} \cap \{3, 4\} \cap \{3, 4\} \cap \cdots = \{3\},$$

---

[30]This follows from the fact that property (4) in Proposition 3 (p. 18) holds with equality when $p = 1$. It follows that the two notions also coincide when B_i^p and B_j^p are replaced with K_i and K_j.

	A	B
A	M,M	$1,-L$
B	$-L,1$	$0,0$

(a) g_a (probability $1-q$)

	A	B
A	$0,0$	$1,-L$
B	$-L,1$	M,M

(b) g_b (probability q)

n_2

	0	1	2	3	4	$\cdots$
0	$1-q$					
1	$q\varepsilon$	$q\varepsilon(1-\varepsilon)$				
2		$q\varepsilon(1-\varepsilon)^2$	$q\varepsilon(1-\varepsilon)^3$			
3			$q\varepsilon(1-\varepsilon)^4$	$q\varepsilon(1-\varepsilon)^5$		
4				$q\varepsilon(1-\varepsilon)^6$	$q\varepsilon(1-\varepsilon)^7$	
$\vdots$					$\ddots$	$\ddots$

(n_1 labels the rows.)

(c) The common prior on the state space.

Figure 3 – The email game again. Payoffs are g_a if the state is $(0,0)$ and g_b otherwise. ε is small, $L > M > 1$ and $q < 1/2$.

and

$$B_j^p(E) \cap B_j^p\left(B_i^p\left(E\right)\right) \cap B_j^p\left(B_i^p\left(B_j^p\left(E\right)\right)\right) \cap \cdots$$
$$= \{1,3,4\} \cap \{1,3,4\} \cap \{1,3,4\} \cap \{1,3,4\} \cap \cdots = \{1,3,4\}.$$

Hence $I^p(E) = \{3\} \cap \{1,3,4\} = \{3\}$.

Morris (1999) argues that I^p is the correct notion of approximate knowledge for some questions (and there are papers that use it). But as we'll see in a moment (section 2.2.3), C^p is the right concept for our purposes (viz. lower hemicontinuity).

Before moving on, let's see what common p-belief has to say about the email game from section 2.1 (p. 14), the essentials of which are replicated in Figure 3. Take any $p > 1/2$ and $\varepsilon > 0$ small,[31] and consider the event $E := \Omega\backslash\{(0,0)\}$ that payoffs are given by g_b rather than g_a. Since player 1 knows the payoffs,

$$B_1^p(E) = \Omega\backslash\{(0,0)\} = E.$$

---

[31] For the argument to follow, a small p will require a smaller ε.

Every $n_2 > 0$ knows E when it occurs, but $n_2 = 0$ believes it only with with probability $q\varepsilon/(1 - q + q\varepsilon)$ when it occurs. Since q and p are $> 1/2$ and ε is small, $n_2 = 0$ does not p-believe E, so

$$B_2^p(E) = E\backslash\{(1,0)\}.$$

Hence

$$E_1 := B_1^p(E) \cap B_2^p(E) = E\backslash\{(1,0)\} = \Omega\backslash\{(0,0),(1,0)\}.$$

Player 2 knows E_1 whenever it occurs, and types $n_1 \geq 2$ of player 1 do as well. But type $n_1 = 1$ of player 1 believes E_1 with probability $1/(2-\varepsilon) < p$ (using the fact that ε is small). Hence

$$E_2 := B_1^p(E_1) \cap B_2^p(E_1) = E_1\backslash\{(1,1)\} = \Omega\backslash\{(0,0),(1,0),(1,1)\}.$$

Continuing inductively, one more state is eliminated in each round. It follows that

$$C^p(E) = \bigcap_{n \in \mathbf{N}} E_n = \varnothing.$$

So in the email game, there is no state at which payoffs are common p-belief for $p > 1/2$. This shows that common p-belief (for p reasonably large) is sometimes a lot stronger than types being close to common-knowledge-of-payoffs types in the product topology. In a word, common p-belief is a demanding condition.

2.2.3 Play under common p-belief is close to play under CK

We are interested in C^p because it is supposed to yield similar behaviour to C in games. The formal results establishing this are from section 5 of Monderer and Samet (1989). We'll just sketch the general idea.

The setup is as follows. Players in a set I with action spaces $\{A_i\}_{i \in I}$ are uncertain about what game they are playing. Formally, there is a measurable space (Ω, Σ) of uncertainty, and payoffs are a random element $G : \Omega \to \mathcal{G}$, where $\mathcal{G}$ is a finite set of complete-information games (with typical element g). Players have information functions $\Pi_i : \Omega \to \Sigma$ and priors μ_i on (Ω, Σ). Putting these ingredients to together gives us a Bayesian game:

$$\gamma := (I, \{A_i\}_{i \in I}, \mathcal{G}, G, (\Omega, \Sigma), \{\Pi_i, \mu_i\}_{i \in I}).$$

Theorem 1 (Monderer and Samet (1989))**.** Take any Bayesian game

$$\gamma = (I, \{A_i\}_{i \in I}, \mathcal{G}, G, (\Omega, \Sigma), \{\Pi_i, \mu_i\}_{i \in I})$$

such that $\mathcal{G}$ is finite and every partition cell of every player has positive probability under each prior,[32] and any collection of Nash equilibria $\{\sigma^g\}_{g \in \mathcal{G}}$. If for some $g \in \mathcal{G}$, $C^p(G^{-1}(g))$ for p sufficiently large has sufficiently high probability according to each player's prior, then there is a nearly-Bayes–Nash equilibrium of γ that nearly always yields the same play and payoffs as σ^G.[33]

In brief, common p-belief of payoffs for p large with high probability is enough for robustness. On the other hand, common p-belief of payoffs for p large with high probability tends to fail when robustness does: for example, we saw in the previous section that in the email game, payoffs are common p-belief for $p > 1/2$ with probability zero.

2.2.4   Perturbations that preserve common p-belief

When studying robustness, we wish to consider a class of perturbations of players' information about payoffs that are 'small' in some sense. We have now seen that if we take 'small deviation from common knowledge of payoffs' to mean 'common p-belief in payoffs for p large with high probability' then robustness is free.

Arguably, this is not a natural (or 'primitive') notion of smallness. More importantly, we already know that it rules out certain perturbations that we want to allow for in our study of robustness, namely the perturbations of g_b in the email game. These perturbations are what motivated the robustness question in the first place: they seem small, and we want to know whether and when they will cause trouble as they do in the email game.

A more primitive notion of 'smallness' of a perturbation of a game in which payoffs are commonly known to be g is to require that players p-believe for p large with high probability that payoffs are g. That is, we require first-order beliefs about payoffs to be close to what they are in the complete-information game g, but leave higher orders of belief unrestricted.

Suppose we consider the class of all perturbations that are 'first-order small' in this sense. When, if ever, do we get the more stringent 'common-p-belief smallness' (and hence play close to CK play) for free? This section provides three sets of conditions.

The first result is a trivial observation: when players' types are independent, mutual p-belief with high probability implies common p-belief with high probability.

---

[32]Formally, every set in $\bigcup_{i \in I} \Pi_i(\Omega) \subseteq \Sigma$ has positive μ_i-measure for each $i \in I$.
[33]Note that σ^G is random since G is.

The second result relies on a finite state space. A proof (and some related results and discussion) can be found in Fudenberg and Tirole (1991, sec. 14.4).

Proposition 5. If Ω is finite, then for any sequence $\{\mu_n\}$ of common priors in $\Delta(\Omega)$ such that $\mu_n(E) \to 1$, there are sequences $\{q_n\}$ and $\{r_n\}$ in $[0, 1]$ converging to 1 such that $\mu_n\left(C^{q_n}(E)\right) \geq r_n$ for every $n \in \mathbf{N}$.

In words, if E is highly likely according to the common prior, then it is highly likely to be common p-belief for p large.

The third result allows the state space to be countably infinite, but requires $p < 1/2$. It is a slight specialisation of Proposition 4.2 in Kajii and Morris (1997b). A proof can be found in their appendix.

Theorem 2 (critical path result)**.** Fix $p \in [0, 1/|I|)$, and assume Ω countable and a common prior μ. Then for any event $E \in \Sigma$ that is the intersection of unions of players' partition cells (a 'simple' event),

$$\mu(C^p(E)) \geq 1 - (1 - \mu(E))\frac{1 - p}{1 - |I|p}.^{34}$$

The crucial corollary is that for fixed p, $\mu\left(C^p(E)\right) \to 1$ as $\mu(E) \to 1$. So if E is highly likely according to the common prior, then it is highly likely to be common p-belief for p large-ish.

The first two results are arguably uninteresting for studying robustness. After all, assuming that types are independent or that there are finitely many states are tacit ways of restricting the information structure. If we're serious about robustness, we should allow for a richer class of perturbations than these conditions permit. The critical path result holds more promise for robustness. In section 2.3.3 (p. 29), we'll see how it can be chained together with a result along the lines of Theorem 1 (p. 22) to prove a positive result about robustness.

An aside: while the first two sets of conditions extend straightforwardly to the case of mutually continuous priors, the critical path result breaks down completely without a common prior. Kajii–Morris robustness (as studied section 2.3.3 below) is therefore extremely demanding in settings without a common prior—see Oyama and Tercieux (2010).

The email game provides a nice way to explore the boundaries of these results. Finiteness of Ω is evidently essential for Proposition 5: as soon as the state space is allowed to be countably infinite, pathologies such as the email

---

[34]Corollary 4.3 in the paper provides a slightly weaker bound for non-simple E.

	A	B
A	M,M	$1,-L$
B	$-L,1$	$0,0$

(a) g_a (probability $1-q$)

	A	B
A	$0,0$	$1,-L$
B	$-L,1$	M,M

(b) g_b (probability q)

n_2

n_1	0	1	2	3	4	$\cdots$
0	$1-q$					
1	$q\varepsilon$	$q\varepsilon(1-\varepsilon)$				
2		$q\varepsilon(1-\varepsilon)^2$	$q\varepsilon(1-\varepsilon)^3$			
3			$q\varepsilon(1-\varepsilon)^4$	$q\varepsilon(1-\varepsilon)^5$		
4				$q\varepsilon(1-\varepsilon)^6$	$q\varepsilon(1-\varepsilon)^7$	
$\vdots$					$\ddots$	$\ddots$

(c) The common prior on the state space.

Figure 4 – The email game again. Payoffs are g_a if the state is $(0,0)$ and g_b otherwise. ε is small, $L > M > 1$ and $q < 1/2$.

game become possible. Recall that we showed in section 2.2 that for $p > 1/2$, it is not common p-belief at any state that payoffs are g_b. The critical path result tells us that $p \geq 1/2$ is essential for such a result: if $p < 1/2$, then for any two-player game (including the email game), a highly likely event must be commonly p-believed with high probability.

The email game (replicated in Figure 4) also provides some intuition for the shape of the bound in the critical path result. Consider again the (clearly simple) event $E = \Omega \backslash \{(0,0)\}$ that payoffs are given by g_b rather than g_a, and fix some $p \in [0,1]$. For convenience, relabel the states according to the order in which we kicked them out in the contagion argument:

$$\omega_0 \coloneqq (0,0), \quad \omega_1 \coloneqq (1,0), \quad \omega_2 \coloneqq (1,1), \quad \omega_3 \coloneqq (2,1), \quad \omega_4 \coloneqq (2,2), \quad \dots$$

By the contagion logic from before, having kicked out states $\omega_0, \dots, \omega_{n-1}$, we kick out state ω_n iff $\mu(\omega_n)/[\mu(\omega_n) + \mu(\omega_{n-1})] < p$, which rearranges to

$$\mu(\omega_n) < \frac{p}{1-p}\mu(\omega_{n-1}) < \left(\frac{p}{1-p}\right)^{n-1}\mu(\omega_0) \quad \text{for each } n \in \mathbf{N}.$$

When we analysed the email game in section 2.1, we assumed $p > 1/2$ and that ε was sufficiently small given p; the contagion argument then

25

knocked out every state. Without these assumptions, the contagion may stop. Suppose that there are at most N rounds of contagion; then we know that $\omega_{N+1}, \omega_{N+2}, \ldots$ are in $C^p(E)$, so

$$
\begin{aligned}
\mu(C^p(E)) &\geq \sum_{n=N+1}^{\infty} \mu(\omega_n) \\
&= 1 - \sum_{n=1}^{N} \mu(\omega_n) \\
&> 1 - \sum_{n=1}^{N} \left(\frac{p}{1-p}\right)^{n-1} \mu(\omega_0) \\
&= 1 - \mu(\omega_0) \sum_{n=0}^{N-1} \left(\frac{p}{1-p}\right)^{n} \\
&= 1 - (1 - \mu(E)) \sum_{n=0}^{N-1} \left(\frac{p}{1-p}\right)^{n}.
\end{aligned}
$$

For N large, if $p \geq 1/2$ then this bound becomes negative, so take $p < 1/2$. Then we have a bound independent of the number of rounds of contagion:

$$
\begin{aligned}
\mu(C^p(E)) &\geq 1 - (1 - \mu(E)) \sum_{n=0}^{\infty} \left(\frac{p}{1-p}\right)^{n} \\
&= 1 - (1 - \mu(E)) \frac{1-p}{1-2p}.
\end{aligned}
$$

This is the bound in the critical path result (for $|I| = 2$). The critical path result is derived by showing that (a generalised version of) the email game's 'contagion' information structure is the worst case for $\mu(C^p(E))$. (The proof is quite long and involved.)

2.3 Kajii and Morris (1997b)

The email game provides a negative result about robustness: strict Nash equilibria need not be robust. Perhaps the best-known positive robustness results are those of Kajii and Morris (1997b).

2.3.1 Setting

The setup is similar to that in section 2.2.3 (p. 22): players in a set I, action spaces $\{A_i\}_{i \in I}$, a set $\mathcal{G}$ of complete-information games, a measurable space

of uncertainty (Ω, Σ), random payoffs $G : \Omega \to \mathcal{G}$, information functions $\Pi_i : \Omega \to \Sigma$, and priors μ_i on (Ω, Σ). Together, these ingredients make up a Bayesian game:

$$\gamma := (I, \{A_i\}_{i \in I}, \mathcal{G}, G, (\Omega, \Sigma), \{\Pi_i, \mu_i\}_{i \in I}).$$

Write $u_i : A \times \mathcal{G} \to \mathbf{R}$ for player i's payoff function.

Assume a common prior μ and that the state space Ω is countable. It is then wlog to let μ have full support, so that conditional probabilities are unique. It is uniqueness of conditional probabilites we're after—we ought to be able to recover all the results (modulo measure-theoretic issues) with Ω uncountable. By contrast, the common-prior assumption is crucial.

On the other hand, we no longer need to consider only a finite set $\mathcal{G}$ of games. On the contrary, we will consider a very rich (definitely uncountable) class of perturbations of the complete-information game g for which we wish to make robust predictions.

An action distribution $\nu \in \Delta(A)$ is an η-correlated equilibrium of a complete-information game g iff

$$\sum_{(a_i, a_{-i}) \in A} u_i(a_i, a_{-i}, g)\nu(a_i, a_{-i}) \geq \sum_{(a_i, a_{-i}) \in A} u_i(\phi(a_i), a_{-i}, g)\nu(a_i, a_{-i}) - \eta$$

for every $\phi : A_i \to A_i$, for every $i \in I$. For $\eta = 0$, we say 'correlated equilibrium' (simpliciter). A Nash equilibrium is a correlated equilibrium ν such that

$$\nu(a) = \prod_{i \in I} \nu_i(a_i) \quad \forall a \in A$$

for some collection of distributions $\{\nu_i\}_{i \in I}$ on $\{A_i\}_{i \in I}$.

A (mixed) strategy for i in the Bayesian game γ is a Π_i-measurable map $\sigma_i : \Omega \to \Delta(A_i)$. A strategy profile is a map $\sigma : \Omega \to \Delta(A)$ whose marginals $\{\sigma_i\}_{i \in I}$ are strategies. Extend u_i to mixed strategies in the usual way. A strategy profile σ is a Bayes–Nash equilibrium (BNE) of γ iff

$$\sum_{\omega \in \Omega} u_i(\sigma_i(\omega), \sigma_{-i}(\omega), G(\omega))\mu(\omega) \geq \sum_{\omega \in \Omega} u_i(\phi(\omega), \sigma_{-i}(\omega), G(\omega))\mu(\omega)$$

for any Π_i-measurable $\phi : \Omega \to A_i$, for every $i \in I$. $\nu \in \Delta(A)$ is a BNE action distribution for γ iff some BNE σ of γ induces it according to

$$\nu(\cdot) = \sum_{\omega \in \Omega} \sigma(\omega)(\cdot)\mu(\omega).$$

For $\varepsilon \in [0,1]$, the Bayesian game γ is an ε-elaboration of the complete-information game g iff $\mu(G^{-1}(g)) = 1 - \varepsilon$, i.e. payoffs are given by g with probability $1 - \varepsilon$. Let $\Gamma(g, \varepsilon)$ denote the set of all ε-elaborations of g. (A very large set.)

g is obviously a 0-elaboration of itself; we might call it the degenerate 0-elaboration. But there are also 0-elaborations with other information structures which allow players to correlate their actions. In fact, an action distribution $\nu \in \Delta(A)$ is a BNE action distribution of some $\gamma \in \Gamma(g, 0)$ iff it is a correlated equilibrium of g (Aumann, 1987).

Endow $\Delta(A)$ with the uniform metric:

$$\|\nu - \nu'\| := \max_{a \in A} |\nu(a) - \nu'(a)| \quad \text{for } \nu, \nu' \in \Delta(A).$$

'ν is δ-close to ν'' will serve as a synonym for '$\|\nu - \nu'\| \leq \delta$'. We will measure the distance between strategy profiles in terms of the distance between the action distributions they induce. It is important to note this is an ex-ante (or 'on-average') notion of closeness: two very different strategy profiles can induce the same action distribution.

Definition 3 (Kajii–Morris robustness)**.** An action distribution $\nu \in \Delta(A)$ is KM-robust to incomplete information in the complete-information game g iff for every $\delta > 0$, there is $\bar{\varepsilon} > 0$ such that every $\gamma \in \bigcup_{\varepsilon \in [0, \bar{\varepsilon}]} \Gamma(g, \varepsilon)$ has a BNE action distribution δ-close to ν.

Since $\Gamma(g, 0)$ contains the degenerate 0-elaboration (i.e. g itself), any robust action distribution must be a Nash (hence correlated) equilibrium. We can therefore speak of robust equilibria rather than of robust action distributions. The converse is not true: there are open sets of games that have strict Nash equilibria but no robust equilibria (see section 3.1 of the paper).

KM-robustness is a lower hemicontinuity condition: for any equilibrium σ of g, any nearby game (=elaboration) must have an equilibrium close to σ. (Note the order of the quanitifiers: $\forall/\exists$.) We take 'nearby game' to mean 'ε-elaboration for ε small', and 'nearby equilibrium' to mean 'action distributions close in the uniform metric'.

2.3.2 Unique correlated equilibria are KM-robust

Kajii and Morris's (1997b) first positive robustness result is as follows.

Proposition 6 (unique CE)**.** If ν is the unique correlated equilibrium of g, then ν is KM-robust in g.

The usefulness of this result is limited by the fact that few games have a unique correlated equilibrium. The authors list some stringent sufficient conditions for a game to have a unique correlated equilibrium on p. 1294.

The proof is pretty easy. Recall Aumann's (1987) result that ν is the action distribution induced by a BNE of some $\gamma \in \Gamma(g, 0)$ iff it is a correlated equilibrium. A small modification to the (straightforward) proof of the 'only if' direction of Aumann's result yields the following:

Lemma 1. For any g and $\eta > 0$, there is $\bar{\varepsilon} > 0$ such that every BNE action distribution of every $\gamma \in \bigcup_{\varepsilon \in [0, \bar{\varepsilon}]} \Gamma(g, \varepsilon)$ is an η-correlated equilibrium of g.

Corollary 1. Suppose that ν_n is a BNE action distribution of $\gamma_n \in \Gamma(g, \varepsilon_n)$ for each $n \in \mathbf{N}$ and that $\varepsilon_n \longrightarrow 0$. Then there is a subsequence of $\{\nu_n\}$ that converges to a correlated equilibrium of g.

Proof of Proposition 6. Suppose that ν is the unique correlated equilibrium of g, and that it is not KM-robust. Then there is a sequence $\{\gamma_n\}$ of elaborations $\gamma_n \in \Gamma(g, \varepsilon_n)$ with $\varepsilon_n \longrightarrow 0$ such that for any sequence of BNE action distributions $\{\nu_n\}$ of $\{\gamma_n\}$, $\{\|\nu_n - \nu\|\}$ is bounded away from zero. On the other hand, $\{\nu_n\}$ has a subsequence converging to ν by Corollary 1. Contradiction! ∎

2.3.3 p-dominant equilibria (for p small) are KM-robust

Kajii and Morris's (1997b) second positive robustness result makes use of the line of reasoning suggested in earlier sections. First, a result along the lines of Theorem 1 (p. 22) tells us that if payoffs are common p-belief with high probability, then lower hemicontinuity obtains. Second, the critical path result (p. 24) tells us that for Ω countable and $p < 1/|I|$, p-belief of payoffs with high probability implies common p-belief of payoffs with high probability.

An action profile $a \in A$ is p-dominant in g iff for every $i \in I$ and every $\lambda \in \Delta(A_{-i})$ with $\lambda(a_{-i}) \geq p$,

$$\sum_{a_{-i} \in A_{-i}} u_i(a_i, a_{-i}, g)\lambda(a_{-i}) \geq \sum_{a_{-i} \in A_{-i}} u_i(a_i', a_{-i}, g)\lambda(a_{-i}) \quad \forall a_i' \in A_i.$$

Dominant-strategy equilibria are 0-dominant, and pure-strategy Nash equilibria are 1-dominant. Any strict Nash equilibrium is p-dominant for some $p < 1$. So a Nash equilibrium is 'more dominant' if its p is lower.

The requisite 'result along the lines of Theorem 1' is the following. The proof is easy—see the pp. 1297–8 of the paper.

Lemma 2. Suppose that $a \in A$ is p-dominant in g, that $\gamma \in \Gamma(g, \varepsilon)$ for some $\varepsilon > 0$, and that $G^{-1}(g)$ is commonly p-believed at some state $\omega \in \Omega$. Then γ has a BNE σ such that $\sigma(\cdot)(a) = 1$.

Proposition 7 (p-dominant)**.** If a is a p-dominant equilibrium of g for $p < 1/|I|$, then the action distribution degenerate at a is KM-robust in g.

Proof. Let a be p-dominant in g for $p < 1/|I|$, and write ν^a for the action distribution degenerate at a. Fix $\delta > 0$. By the critical path result (p. 24), we can find $\varepsilon > 0$ such that for any

$$\gamma := (I, \{A_i\}_{i \in I}, \mathcal{G}, G, (\Omega, \Sigma), \{\Pi_i, \mu_i\}_{i \in I})$$

in $\Gamma(g, \varepsilon)$, $\mathbf{P}(G^{-1}(g)) > 1 - \varepsilon$ implies $\mathbf{P}(C^p(G^{-1}(g))) > 1 - \delta$. It follows by Lemma 2 that γ has a BNE σ such that $\sigma(\cdot)(a) = 1$, and hence that γ has a BNE action distribution ν s.t. $\nu(a) \geq \mathbf{P}(C^p(G^{-1}(g))) > 1 - \delta$. So for any $\delta > 0$, there's a BNE action distribution ν s.t. $\|\nu - \nu^a\| < \delta$. ∎

A corollary is that in 2×2 games, a strictly risk-dominant equilibrium is KM-robust. This substantially strengthens Carlsson and van Damme's (1993) global-games/'robustness' argument for risk dominance. (Because their result only concerns robustness to a very small class of elaborations.)

2.3.4 Extensions

Oyama and Tercieux (2010) study a strengthening of KM-robustness that allows for elaborations without a common prior. They show that Kajii and Morris's (1997b) positive results about robustness break down spectacularly: for generic games, a pure-strategy equilibrium is robust in this sense iff it is uniquely rationalisable![35]

Kajii and Morris's (1997b) sufficient conditions for robustness in general games are fairly stringent. By limiting attention to a smaller class of games, we might hope to find tighter sufficient conditions for robustness. In this vein, Ui (2001) shows that the potential-maximiser in a cardinal potential game is always robust. Morris and Ui (2005) extend his result to ordinal potential games. There are other examples.

Kajii and Morris's (1997b) results are for games in normal form. We might also be interested in what robustness demands in extensive-form games. Chassang and Takahashi (2011) extend the critical path result (p. 24) to

---

[35]This is obviously related to Weinstein and Yıldız (2007), discussed below.

repeated games, and use it to develop a recursive characterisation of KM-robustness in repeated games. (And then prove a robust folk theorem, and study robust cooperation in the repeated prisoners' dilemma).[36]

2.4 Robustness of refinements

KM-robustness is just lower hemicontinuity: for any equilibrium σ of g, any nearby game (=elaboration) must have an equilibrium close to σ. The two quantifiers are, respectively, '$\forall$ nearby games' and '$\exists$ nearby equilibrium'. We can ask other robustness-type questions by fiddling with the quantifiers. '$\forall$ nearby games'/'$\forall$ nearby equilibria' is obviously too demanding a notion of robustness (due to equilibrium multiplicity), so discard that one.

Robustness notions with $\exists$ as the first quantifier are too weak since every game is nearby itself. One way to obtain a stronger robustness concept is to replace '$\exists$ nearby game' with '$\exists$ *interesting* nearby game'. In particular, we require there to be a nearby game that satisfies a payoff richness condition to the effect that every action is strictly dominant for some type / with positive probability.

The idea is that if an equilibrium can be 'justified' by some nearby payoff-rich games, then an analyst with imperfect information about payoffs cannot rule out that it will be played. Such an equilibrium is 'robust to refinement': it cannot credibly be refined away. If we obtain a positive result about some notion of robustness of this type, then we can conclude that any refinement that rules it out is a 'non-robust' refinement.

The two main papers with refinement-killing results of this type are Dekel and Fudenberg (1990) and Weinstein and Yıldız (2007).[37] The former use a notion of robustness with '$\exists$ nearby equilibrium', whereas the latter use '$\forall$ nearby equilibrium'. An additional important difference is that they use different notion of closeness of games. Dekel and Fudenberg (1990) (like Kajii and Morris (1997b)) use an ex-ante notion of closeness of games: two games are close iff payoffs are close with high prior probability. Weinstein and Yıldız (2007) take the 'interim perspective' (i.e. they eschew priors entirely), and so measure closeness of games according to closeness in the product topology on the universal type space.[38]

---

[36]Unsurprisingly, there are some technical issues that have to be dealt with. One of them is that $G^{-1}(g)$ needs to be redefined slightly.

[37]As in Kajii and Morris (1997b), the positive robustness results in both of these papers are arrived at via a critical path result that is derived using a contagion argument.

[38]Although their notion of closeness does not involve prior probabilities, Weinstein and Yıldız (2007) show that their results hold even if we restrict attention to types that are consistent with a common prior.

Dekel and Fudenberg (1990) show that an action profile is refinement-robust in their sense iff it lies in $S^\infty W$, the set of actions surviving one round of deletion of weakly dominated strategies followed by ordinary iterated deletion of strictly dominated strategies.[39] They conclude that any refinement of $S^\infty W$ is non-robust; hence in particular, any refinement of rationalisability (S^∞) is non-robust.[40]

Modulo the difference in how closeness of games is measured, Weinstein and Yıldız's (2007) notion of robustness is stronger than Dekel and Fudenberg's (1990). They correspondingly obtain a slightly weaker result: that any rationalisable action profile is refinement-robust in their sense, so any refinement of rationalisability is non-robust.

Although both results show that refinements of rationalisability are non-robust in some sense, the latter is far more problematic for refinements. If a refinement is non-robust in Dekel and Fudenberg's (1990) sense, then there are games for which it rules out equilibria that can be justified by some equilibrium selection from some payoff-rich class of nearby games. A believer in refinements may respond that there may very well be another equilibrium selection that is far away. There is no such escape from Weinstein and Yıldız (2007): their result says that every rationalisable profile in any game is *uniquely* rationalisable in a nearby game.

---

[39]Börgers (1994) similarly shows that the set of actions consistent with common p-belief in rationality shrinks to $S^\infty W$ as $p \rightarrow 1$. The idea is similar: when players entertain small doubts about other players' rationality, weakly dominated actions are ruled out. So $S^\infty W$, a concept developed in the robustness literature, has epistemic relevance.

[40]This result has an antecedent in Fudenberg, Kreps and Levine (1988), who showed that refinements of Nash equilibrium are non-robust.

References

Ahn, D. S. (2007). Hierarchies of ambiguous beliefs. *Journal of Economic Theory*, *136*(1), 286–301.

Armbruster, W. & Böge, W. (1979). Bayesian game theory. In O. Moeschlin & D. Pallaschke (Eds.), *Game theory and related topics: Proceedings of the seminar on game theory and related topics, Bonn/Hagen, 26–29 September, 1978* (pp. 17–28). Amsterdam: North-Holland.

Aumann, R. J. (1976). Agreeing to disagree. *Annals of Statistics*, *4*, 1236–1239.

Aumann, R. J. (1987). Correlated equilibrium as an expression of Bayesian rationality. *Econometrica*, *55*(1), 1–18.

Aumann, R. J. (1999). Interactive epistemology I: Knowledge. *International Journal of Game Theory*, *28*(3), 263–300.

Battigalli, P. & Siniscalchi, M. (1999). Hierarchies of conditional beliefs and interactive epistemology in dynamic games. *Journal of Economic Theory*, *88*(1), 188–230.

Billingsley, P. (1995). *Probability and measure* (3rd). New York, NY: Wiley.

Billingsley, P. (1999). *Convergence of probability measures* (2nd). New York, NY: Wiley.

Böge, W. & Eisele, T. (1979). On solutions of Bayesian games. *International Journal of Game Theory*, *8*(4), 193–215.

Börgers, T. (1994). Weak dominance and approximate common knowledge. *Journal of Economic Theory*, *64*(1), 265–276.

Brandenburger, A. & Dekel, E. (1993). Hierarchies of beliefs and common knowledge. *Journal of Economic Theory*, *59*(1), 189–198.

Carlsson, H. & van Damme, E. (1993). Global games and equilibrium selection. *Econometrica*, *61*(5), 989–1018.

Chassang, S. & Takahashi, S. (2011). Robustness to incomplete information in repeated games. *Theoretical Economics*, *6*(1), 49–93.

Chen, Y.-C. (2010). Universality of the Epstein-Wang type structure. *Games and Economic Behavior*, *68*(1), 389–402.

Dekel, E. & Fudenberg, D. (1990). Rational behavior with payoff uncertainty. *Journal of Economic Theory*, *52*(2), 243–267.

Dekel, E., Fudenberg, D., & Morris, S. (2006). Topologies on types. *Theoretical Economics*, *1*(3), 275–309.

Epstein, L. G. & Peters, M. (1999). A revelation principle for competing mechanisms. *Journal of Economic Theory*, *88*(1), 119–160.

Epstein, L. G. & Wang, T. (1996). 'Beliefs about beliefs' without probabilities. *Econometrica*, *64*(6), 1343–1373.

Epstein, L. G. & Zin, S. E. (1989). Substitution, risk aversion, and the temporal behavior of consumption and asset returns: A theoretical framework. *Econometrica*, *57*(4), 937–969.

Fudenberg, D., Kreps, D. M., & Levine, D. K. (1988). On the robustness of equilibrium refinements. *Journal of Economic Theory*, *44*(2), 354–380.

Fudenberg, D. & Tirole, J. (1991). *Game theory*. Cambridge, MA: MIT Press.

Gül, F. & Pesendorfer, W. (2001). Temptation and self-control. *Econometrica*, *69*(6), 1403–1435.

Gül, F. & Pesendorfer, W. (2016). Interdependent preference models as a theory of intentions. *Journal of Economic Theory*, *165*, 179–208.

Harsányi, J. C. (1967). Games with incomplete information played by 'Bayesian' players, I–III: Part I. The basic model. *Management Science*, *14*(3), 159–182.

Heifetz, A. (1993). The Bayesian formulation of incomplete information: The non-compact case. *International Journal of Game Theory*, *21*(4), 329–338.

Heifetz, A. & Samet, D. (1998). Topology-free typology of beliefs. *Journal of Economic Theory*, *82*(2), 324–341.

Heifetz, A. & Samet, D. (1999). Coherent beliefs are not always types. *Journal of Mathematical Economics*, *32*(4), 475–488.

Kajii, A. & Morris, S. (1997a). Common *p*-belief: The general case. *Games and Economic Behavior*, *18*(1), 73–82.

Kajii, A. & Morris, S. (1997b). The robustness of equilibria to incomplete information. *Econometrica*, *65*(6), 1283–1309.

Kohlberg, E. & Mertens, J.-F. (1986). On the strategic stability of equilibria. *Econometrica*, *54*(5), 1003–1037.

Kreps, D. M. & Porteus, E. L. (1978). Temporal resolution of uncertainty and dynamic choice theory. *Econometrica*, *46*(1), 185–200.

Lipman, B. L. (1991). How to decide how to decide how to... Modeling limited rationality. *Econometrica*, *59*(4), 1105–1125.

Mariotti, T., Meier, M., & Piccione, M. (2005). Hierarchies of beliefs for compact possibility models. *Journal of Monetary Economics*, *41*(3), 303–324.

Meier, M. (2012). An infinitary probability logic for type spaces. *Israel Journal of Mathematics*, *192*(1), 1–58.

Mertens, J.-F., Sorin, S., & Zamir, S. (2015). *Repeated games*. Cambridge: Cambridge University Press.

Mertens, J.-F. & Zamir, S. (1985). Formulation of Bayesian analysis for games with incomplete information. *International Journal of Game Theory*, *14*(1), 1–29.

Milgrom, P. (1981). An axiomatic characterization of common knowledge. *Econometrica*, *49*(1), 219–222.

Monderer, D. & Samet, D. (1989). Approximating common knowledge with common beliefs. *Games and Economic Behavior*, *1*(2), 170–190.

Morris, S. (1999). Approximate common knowledge revisited. *International Journal of Game Theory*, *28*(3), 385–408.

Morris, S. & Ui, T. (2005). Generalized potentials and robust sets of equilibria. *Journal of Economic Theory*, *124*(1), 45–78.

Myerson, R. B. (1986). Multistage games with communication. *Econometrica*, *54*(2), 323–358.

Myerson, R. B. (1991). *Game theory: Analysis of conflict*. Cambridge, MA: Harvard University Press.

Osborne, M. J. & Rubinstein, A. (1994). *A course in game theory*. Cambridge, MA: MIT Press.

Oyama, D. & Tercieux, O. (2010). Robust equilibria under non-common priors. *Journal of Economic Theory*, *145*(2), 752–784.

Rényi, A. (1955). On a new axiomatic theory of probability. *Acta Mathematica Academiae Scientiarum Hungarica*, *6*(3), 285–335.

Rubinstein, A. (1989). The electronic mail game: Strategic behavior under 'almost common knowledge'. *American Economic Review*, *79*(3), 385–391.

Siniscalchi, M. (2008). Epistemic game theory: Beliefs and types. In S. N. Durlauf & L. E. Blume (Eds.), *The new Palgrave dictionary of economics* (Vol. 2, pp. 895–899). London: Palgrave Macmillan.

Ui, T. (2001). Robust equilibria of potential games. *Econometrica*, *69*(5), 1373–1380.

Weinstein, J. & Yıldız, M. (2007). A structure theorem for rationalizability with application to robust predictions of refinements. *Econometrica*, *75*(2), 365–400.