



University  
of Glasgow

Adam Smith  
Business School

# WORKING PAPER SERIES



Partial Consensus in Large Games and Markets

Gabriel Desgranges and Sayantan Ghosal

Paper no. 2021-02  
February 2021

# Partial Consensus in Large Games and Markets\*

GABRIEL DESGRANGES AND SAYANTAN GHOSAL  
*Université de Cergy-Pontoise and University of Glasgow*

February 2, 2021

## Abstract

When is partial consensus compatible with equilibrium and when does it lead to non-equilibrium outcomes in large games and markets? In this paper, (a) we develop a new solution concept that allows for partial consensus about the outcomes of strategic and market interaction, and (b) an associated, continuous measure of the degree of stability, via belief coordination, for equilibrium outcomes. We differentiate the properties of our concepts from related notions developed elsewhere. In an economic application we examine the foundations of intertemporal trade via belief coordination in a two period economy and show that, under certain conditions, partial consensus over future prices is consistent with an asset price bubble.

Keywords:  $p$ -consensus,  $p$ -stability, common knowledge, rationalizability, heterogeneous beliefs, coordination, games, markets.

JEL Classifications C70 and D84.

---

\*We would like to thank seminar/conference audiences at Warwick, Cergy-Pontoise, Barcelona, Atlanta for their helpful comments. Both authors would like to acknowledge funding from Université Cergy-Pontoise and University of Glasgow; Ghosal is grateful for funding from the ESRC funded Rebuilding Macro Network. E-mail: gabriel.desgranges@u-cergy.fr, Sayantan.Ghosal@glasgow.ac.uk.

# 1 Introduction

When is partial consensus compatible with equilibrium and when does it lead to non-equilibrium outcomes in large games and markets? How does the possibility that partial consensus is "approximately" self-fulfilling change the analysis of strategic interaction and market behavior? What additional economic insights can be obtained?

In this paper, we take a first step toward answering these questions by developing (i) a new solution concept that allows for partial consensus about the outcomes of strategic and market interaction, (ii) an associated, continuous measure of the degree of stability, via belief coordination, for equilibrium outcomes, (iii) examples, illustrating and differentiating the properties of these two concepts from related notions developed elsewhere, and (iv) an economic application examining the foundations of intertemporal trade via belief coordination in a two period economy where, under certain conditions, partial consensus over future prices is consistent with an asset price bubble.

Our starting point is that agents' heterogeneous beliefs about the outcomes of strategic and market interaction may be characterized by partial consensus. We model partial consensus as common knowledge of  $p$ -beliefs (beliefs that put probability at least  $p$  on a specific outcome); with probability  $1 - p$  beliefs over outcomes are heterogeneous and aren't required to be common knowledge. We model belief coordination with partial consensus and for an outcome to be "approximately" self-fulfilling, we require it to be consistent with all that is commonly known by each agent via an iterative elimination process. Heuristically, our solution concept for large games and markets,  $p$ -consensus, requires that an outcome, about which there is partial consensus, is "approximately self-fulfilling" as referred to in the preceding sentence. Associated with our solution concept is a continuous index of stability,  $p$ -stability, which requires equilibrium to be the unique rationalizable outcome compatible with  $p$ -beliefs. The continuous index for stability we develop can be assigned to each equilibrium within a large class of games. It is less coarse than the usual "stable/unstable" typology and justifies, in smooth applications, the intuition that the stability of equilibrium relates to the slope of the best response map.

The two concepts introduced by us bring added value to the analysis of large games and markets only when there are multiple rationalizable outcomes. Our results explore this point in greater detail.

We begin by demonstrating the link between the set of rationalizable outcomes and the set of  $p$ -consensus outcomes: any outcome in the interior of the set of rationalizable outcomes is also  $p$ -consensus outcome for some  $p > 0$ , thus characterizing the set of outcomes over which partial consensus can be "approximately" self-fulfilling (with  $p$  serving as a measure of the approximation). We say that a Nash equilibrium is inadmissible if the best-response map is "vertical at the equilibrium" i.e. a small change in the other players actions implies an infinitely large change in any one player's best response. We show that every admissible Nash equilibrium is  $p$ -stable for some  $p < 1$ . Under an additional, mild continuity restriction, we show that a  $p$ -stable equilibrium is

a locally isolated  $p$ -consensus distribution i.e. there is no other  $p$ -consensus distribution in its vicinity. Under an additional interiority condition, we also prove the converse statement: if an equilibrium is not  $p$ -stable, then there is always a  $p$ -consensus outcome in its vicinity.

Next, using a number of examples, where the best responses of players depend only on the average action, we work obtain closed form solutions for  $p$ -consensus outcomes and characterize the  $p$ -stability of equilibrium outcomes. We check that these can be supported as the outcomes of a correlated equilibrium with such two states, and in the case of strategic complements (but not strategic substitutes), are consistent with a common prior. In smooth settings, setting we show that the degree of stability is related to the inverse of the slope of the best-response, thus confirming the intuition that the slope of the best-response matters for the strategic stability of equilibrium. In a setting with multiple equilibria, we examine the link between the size of the basin of attraction of an equilibrium outcome and  $p$ -stability and we show that two aren't related.

Our results generalize, the intuition obtained in a game with a finite number of pure strategies that any strict Nash equilibrium is  $p$ -dominant (Morris, Rob and Shin (1995)). We show that the approximate common of beliefs (and the corresponding concept of belief potential) is quite different from the common knowledge of  $p$ -beliefs (which is needed for the definition of  $p$ -consensus and  $p$ -stability) because approximate common knowledge of a Nash equilibrium profile of actions is consistent with the fact that the Nash equilibrium profile of actions is not  $p$ -believed anywhere (while this is required for the definition of  $p$ -consensus and  $p$ -stability). We examine the link between  $p$ -BR, iterated  $p$ -BR (Tercieux, 2006) and  $p$ -stability. We show that although iterated  $p$ -BR sets and  $p$ -stability are equivalent in finite games,  $p$ -stability is adapted to game with continuous set of actions, while iterated  $p$ -stability is not. In economic analysis of large games and markets the typical assumption is that agents have continuous action sets: hence, the concepts,  $p$ -consensus and  $p$ -stability, proposed here enable the analysis of the implications of partial consensus in such settings.

Next, we examine the foundations of intertemporal trade via belief coordination in a two period economy. In contrast to Chatterji and Ghosal (2004), we assume that individuals submit demand functions so that belief coordination in current spot market prices is never a problem: instead the individuals need to coordinate beliefs about future prices when they trade in current spot markets. We show that when redistributions of revenue in second period markets change or when redistributions of commodities in period 1 change second period spot market prices *and* there is lack of consensus over second period prices (so that beliefs over second period prices are heterogeneous to a sufficient degree), then an asset price bubble exists.

The notion of  $p$ -consensus and  $p$ -stability builds on the seminal work on eductive stability (Guesnerie (1992), Evans and Guesnerie (1993), Guesnerie (2005), Evans, Guesnerie and McCulloch (2019)). The analysis developed here differs in that we propose a new solution concept that allows belief heterogene-

ity and our measure of stability is continuous one. Related solution concepts allowing for heterogeneous beliefs off the equilibrium path of play (e.g. the notion of self-confirming equilibrium, Fudenberg and Levine (1993)) and in the macro literature by Angeletos, Collard and Dellas (2018), Woodford (2013)): in the second paper, belief heterogeneity is treated as part of the data of the economy (in contrast, (iterated) belief coordination is central to our analysis) while the third paper carries out an analysis similar to local eductive stability in New Keynesian model. Although the formal definitions are very different, our solution concept is motivated by concerns similar to the literature on sunspot equilibria (Cass and Shell (1983), Farmer and Platonov (2019)).

The next section introduces the formal model within which we define our two concepts, provides a number of basic results. Section 3 illustrates the workings of the two concepts using a number of examples. Section 4 is devoted to finite games. In section 5, we study partial consensus in a two period economy. The last section concludes. The appendix contains proofs of most of our results and the detailed working out of examples.

## 2 The concepts and basic properties

In this section, we begin by describing the underlying strategic framework. We then define  $p$ -stability and state the conditions under which a Nash equilibrium is  $p$ -stable.

### 2.1 The model

The underlying strategic framework is due to MasColell (1984). Let  $A$  be a non-empty compact metric space of actions and  $\Delta(A)$  be the compact and metrizable set of Borel probability measures on  $A$  endowed with the weak convergence topology which is metrizable using the Prohorov metric. For later reference, we note that the Prohorov metric is defined by:

$$d_{\Delta(A)}(m, \tau_a^*) = \inf \{ \varepsilon > 0 : m(M) \leq \tau_a^*(M^\varepsilon) + \varepsilon \text{ for all Borel subsets } M \text{ of } A \},$$

where  $M^\varepsilon = \{y \in A / d_A(x, y) \leq \varepsilon \text{ for some } x \in M\}$ .<sup>1</sup> Let  $U_A$  be the set of continuous utility functions  $u : A \times \Delta(A) \rightarrow R$  endowed with the supremum norm, the metric, separable and complete space of player characteristics. A game with a continuum of players is a Borel measure  $\mu$  on  $U_A$ . For any probability measure  $\tau$  on  $U_A \times A$ , let  $\tau_u$  and  $\tau_a$  denote the respective marginal distribution on  $U_A$  and  $A$  respectively. Let  $T$  denote the set of probability measures on  $U_A \times A$  such that  $\tau_u = \mu$ ,  $\tau \in T$ :  $T$  denotes the set of "strategy profiles", *i.e.*, a distribution of actions for each  $u$ . For  $\tau \in T$ , let  $B_\tau = \{(u, a) : u(a, \tau_a) \geq u(A, \tau_a)\}$ . The best-response correspondence is a map

<sup>1</sup>See Dudley (1989) for this definition and other properties of the Prohorov metric not explicitly mentioned in this paper.

$\phi : T \rightarrow T$  such that  $\phi(\tau) = \{\tau' \in T : \tau'(B_\tau) = 1\}$ , *i.e.*,  $\phi$  is the set of "strategy profiles" putting probability one on the fact that each player plays a best-response to  $\tau$ . A Nash equilibrium is a measure  $\tau^* \in T$  such that  $\tau^* \in \phi(\tau^*)$ .

**Existence result (Theorem 1 in MasColell (1984)).** For a given  $\mu$ , there exists a Nash equilibrium distribution  $\tau^*$ .

**Remark:** The above framework requires all agents who have the same utility function must also choose the same actions and/or have the same the beliefs over distributions. On the face of it, in an example where all agents have the same utility functions, this would require agents to have "homogeneous" beliefs. But, by re interpreting the framework so that utility functions are differ up to an additive constant, we can use the above framework in setting where agents with the same utility functions have heterogeneous beliefs.

## 2.2 p-consensus and p-stability: definition

For a fixed  $\tau$  and  $p \in [0, 1]$ , a  $p$ -belief is a probability distribution  $\tau_p = p\tau + (1-p)\tau'$  for any  $\tau' \in T$ , *i.e.*, a belief that assigns a probability  $p$  to the distribution  $\tau$ . Let  $T_{\tau,p} \subseteq T$  denote the corresponding set.

We define a  $p$ -consensus distribution iteratively as follows. Let  $S_{\tau,p}^0 = T_{\tau,p}$  and consider the sequence of sets  $S_{\tau,p}^n(\tau) = [\phi(S_{\tau,p}^{n-1})] \cap T_{\tau,p}$  for  $n \geq 1$ . This sequence is decreasing and therefore, it converges to a set  $S_{\tau,p}^\infty$ . Then,  $\tau$  is a  $p$ -consensus distribution if  $\tau \in S_{\tau,p}^\infty$ .

For any  $p' < p$ ,  $p, p' \in [0, 1]$ , we have that  $S_{\tau,p}^n \subseteq S_{\tau,p'}^n$  as  $T_{\tau,p} \subseteq T_{\tau,p'}$ . Therefore,  $S_{\tau,p}^\infty \subseteq S_{\tau,p'}^\infty$  and if  $\tau \in S_{\tau,p}^\infty$  then  $\tau \in S_{\tau,p'}^\infty$ . It follows only a rationalizable distribution can be a  $p$ -consensus distribution and for a rationalizable distribution, the set  $I_\tau = \{p \in [0, 1] : \tau \in S_{\tau,p}^\infty\}$  is either empty or  $[0, p]$  for some  $p \in [0, 1]$ .

**Definition 1.** When  $I_\tau$  is non-empty, we say  $\tau$  is a  $p$ -consensus distribution for  $p = \sup I_\tau$ .

Note that a standard definition of rationalizability in this set-up would be to require  $\tau \in S_{\tau,0}^\infty$  (note that this set doesn't depend on the choice of  $\tau$ ). It follows that  $I_\tau$  is non-empty iff  $\tau$  is rationalizable. In this sense,  $p$ -consensus is a refinement of rationalizability and the interesting question is whether  $\tau$  is a  $p$ -consensus distribution for  $p \neq 0$ .

When  $p = 1$ , the set of 1-consensus distributions and Nash equilibrium distributions must coincide.

We define a Nash equilibrium  $\tau^*$  to be  $p$ -stable if the equilibrium distribution is the only element surviving the iterated elimination of non best-responses to a  $p'$ -belief for all  $p' > p$ . This definition relies on a "standard" definition of rationalizable outcomes in a game where the strategy set is restricted to  $T_{\tau^*,p}$ : a  $p$ -stable equilibrium is an equilibrium that is the only rationalizable outcome in a game with the restricted strategy set  $T_{\tau^*,p}$ .

For any  $p < p'$ ,  $p, p' \in [0, 1]$ , note that  $S_{\tau^*,p'}^\infty \subseteq S_{\tau^*,p}^\infty$  and if  $S_{\tau^*,p}^\infty = \{\tau^*\}$  then  $S_{\tau^*,p'}^\infty = \{\tau^*\}$ . In particular, the set  $J_{\tau^*} = \{p \in [0, 1] : S_{\tau^*,p}^\infty = \{\tau^*\}\}$  is an interval of the form  $[p, 1]$  for  $p \in [0, 1]$ .

**Definition 2.** A Nash equilibrium  $\tau^*$  is  $p$ -stable if  $p = \inf J_{\tau^*} = \{p' \in [0, 1] : S_{\tau^*, p'}^\infty = \{\tau^*\}\}$ . The interesting question is whether  $p < 1$ .

**Remark:**

1. We do not require that  $S_{\tau^*, p}^\infty = \{\tau^*\}$  for a  $p$ -stable equilibrium. In some classes of games (for example in the smooth one-dimensional case below),  $S_{\tau^*, p}^\infty \neq \{\tau^*\}$  at a  $p$ -stable equilibrium.

2. If  $\tau^*$  is 0-stable then  $\tau^*$  is the unique rationalizable outcome.

3. If  $\tau^*$  is  $p'$ -dominant then it follows that  $S_{\tau^*, p'}^1 = \{\tau^*\}$ . Hence,  $S_{\tau^*, p'}^\infty = \{\tau^*\}$  and therefore,  $\tau^*$  is  $p$ -stable where  $p \leq p'$ .

4. The two definitions above bring added value in the analysis of strategic outcomes only when there are multiple rationalizable distributions. In what follows, this point is explored in greater detail.

### 2.3 Characterization of $p$ -consensus and $p$ -stability

We begin by providing an existence result for  $p$ -consensus distributions with  $p \neq 0$ .

**Proposition 1.** Suppose the set  $S_0^\infty$  of rationalizable distributions has a non-empty interior in  $T$ . Any distribution in the interior of  $S_0^\infty$  is  $p$ -consensus distribution for some  $p \neq 0$ .

**Proof of Proposition 1.** Consider  $\hat{\tau} \in \text{Int}.S_0^\infty$  and a small neighborhood  $N \subset \text{Int}.S_0^\infty$  of  $\hat{\tau}$ . For  $\tau \in N$ , by upper hemi-continuity of  $\phi$ , there exists  $p$  close enough to zero and  $\tau' \in S_0^\infty$  in a neighborhood  $N'$  of  $\hat{\tau}$  such that  $\phi(p\hat{\tau} + (1-p)\tau') = \tau$ . Denote the corresponding set  $B_{\hat{\tau}, \tau, p}$  and let  $B_{\hat{\tau}, p} = \cup_{\tau \in S_0^\infty} B_{\hat{\tau}, \tau, p}$ . Note that  $B_{\hat{\tau}, p} \subseteq N'$ . Let  $T_{\hat{\tau}, p}(B) = \{p\hat{\tau} + (1-p)\tau, \tau \in B\}$ . Note that  $T_{\hat{\tau}, p}(B_{\hat{\tau}, p}) \subseteq \text{Int}.T_{\hat{\tau}, p}$ . Hence, for any  $\tau$  in  $N$ , for each  $N$  small enough,  $T_{\hat{\tau}, p}(B_{\hat{\tau}, p}) \subseteq \text{Int}.T_{\hat{\tau}, p}$ , so that  $T_{\hat{\tau}, p}(B_{\hat{\tau}, p}) = T_{\tau, p}(B_{\tau, p})$ ; moreover, by upper hemi-continuity of  $\phi$ ,  $B_{\tau, p} \subseteq S_0^\infty$  and hence,  $N \subseteq S_0^\infty$ , as required. ■

Next, we examine the conditions under which an equilibrium is  $p$ -stable. To this end, define a best-response correspondence for each  $u \in U_A$  to each  $m \in \Delta(A)$  as  $B(u, m) = \{a \in A : u(a, m) \geq u(A, m)\}$ : an action in  $B(u, m)$  is a best-response for  $u \in U_A$  to some  $m \in \Delta(A)$ . For each  $m \in \Delta(A)$ , consider the set

$$\tilde{U}_A(m) = \left\{ u \in U_A : B(u, m) \text{ is not single-valued or } \limsup_{m' \rightarrow m} \frac{d_A(B(u, m'), B(u, m))}{d_{\Delta(A)}(m', m)} < \infty \right\},$$

where  $d_A$  denotes a distance on  $A$  and  $d_{\Delta(A)}(\cdot, \cdot)$  denotes the Prohorov metric on  $\Delta(A)$ . Consider two Dirac measures  $\delta_x$  and  $\delta_y$ . Then,  $d_{\Delta(A)}(\delta_x, \delta_y) = d_A(x, y)$ . Note that for each  $u \in \tilde{U}_A(m)$ , a small change in  $m$  induces an infinitely large change in best-responses.

Consider a given  $\tau^*$ . For every  $u$ , let

$$k_{u, \tau} = \limsup_{m \rightarrow \tau_a} \frac{d_A(B(u, m), B(u, \tau_a))}{d_{\Delta(A)}(m, \tau_a)},$$

and

$$K_\tau = \sup_{u \in U_A} \text{ess} k_{u,\tau}.$$

$K_\tau$  is the essential upper bound of  $k_{u,\tau}$  w.r.t. measure  $\mu$  (that is: the set of  $u$  such that  $k_{u,\tau} > K_\tau$  has  $\mu$ -measure 0).

**Definition 3.** The equilibrium  $\tau^*$  of a game  $\mu$  is admissible if  $\mu(\tilde{U}_A(\tau_a^*)) = 0$  and

$$K_\tau < +\infty.$$

As a preliminary step, the following lemma<sup>2</sup> summarizes three key properties of the Prohorov metric that will be useful in proving the result below.

**Lemma 1.** (i) Consider  $\tau = p\hat{\tau} + (1-p)\tau'$ . Then,

$$d_{\Delta(A)}(\tau_a, \hat{\tau}_a) \leq (1-p)d_{\Delta(A)}(\tau'_a, \hat{\tau}_a).$$

(ii) Consider a Dirac measure  $\delta_x$  and a distribution  $\tau_a \in \Delta(A)$ . Consider  $S$  the support of  $\tau_a$  (the smallest closed set s.t.  $\tau_a(S) = 1$ ) and  $d = \sup_{y \in S} d_A(x, y)$  ( $d$  is the radius of the smallest ball centered on  $x$  that contains  $S$ )<sup>3</sup>. Then,

$$d_{\Delta(A)}(\delta_x, \tau_a) \leq d.$$

(iii) Consider  $\tau_a \in \Delta(A)$  defined by  $\tau_a = \int \tau_\lambda f(d\lambda)$  where  $f$  is a probability distribution on a set of parameters  $\lambda$ . Consider another distribution  $\nu \in \Delta(A)$ . We have

$$d_{\Delta(A)}(\tau_a, \nu) \leq \sup_{\lambda} \text{ess} d_{\Delta(A)}(\tau_\lambda, \nu).$$

**Proof.** See appendix. ■

We are now in a position to state and prove the following result:

**Proposition 2.** For any admissible equilibrium, there is a  $\hat{p} < 1$  such that the equilibrium is  $\hat{p}$ -stable.

**Proof.** See appendix. ■

Heuristically, the idea underlying the proof is as follows. An equilibrium  $\tau^*$  is  $p$ -stable if the best-response map, restricted to  $p$ -beliefs, generating the sequence of sets  $S_{\tau^*, p}^n$  is a contraction. For  $p$  close to one, when the equilibrium  $\tau^*$  is admissible, we show that the best-response map, restricted to  $p$ -beliefs, cannot vary much (i.e. in the smooth case, the derivative of the best-response map, restricted to  $p$ -beliefs, is small). If, on the contrary, the equilibrium  $\tau^*$  isn't admissible, even when restricted to  $p$ -beliefs, it can change dramatically around the equilibrium implying that the preceding step of the argument doesn't hold.

We conclude our characterization by proving a result linking  $p$ -consensus and  $p$ -stability.

**Proposition 3.** Consider a  $p$ -stable equilibrium  $\tau^*$ .

<sup>2</sup>We state and prove this lemma for completeness as we are not aware of an explicit proof of the three properties of the Prohorov metric contained in the lemma and these required for the proof of Proposition 1 below.

<sup>3</sup>Notice that  $x$  may be in  $S$  or not.



(i) Suppose  $K_\tau$  is continuous in  $\tau$ . Then, for any  $\hat{p} > p$ , there is a neighborhood of  $\tau^*$  such that no  $\tau$  belonging to the neighborhood is a  $p'$ -consensus distribution for a value  $p' \geq \hat{p}$ .

(ii) Suppose  $S_{\tau^*,p}^\infty$  has a non-empty interior in  $T$  and  $\phi$  is a continuous correspondence. Then, for any  $\hat{p} < p$ , there exists a  $\tau \in S_{\tau^*,p}^\infty$  arbitrarily close to  $\tau^*$  such that  $\tau$  is a  $p'$ -consensus distribution for some  $p' \geq \hat{p}$ .

**Proof.** (i) The proof relies on the proof of Proposition 2: A straightforward necessary condition for  $p$ -stability is  $K_{\tau^*}(1-p) \leq 1$ ; hence,  $K_{\tau^*}(1-\hat{p}) < 1$ ; by the smoothness assumption it follows that for  $\tau$  arbitrarily close to  $\tau^*$ ,  $K_\tau(1-\hat{p}) < 1$ ; so  $S_{\tau,\hat{p}}^\infty$  is either empty or has a radius 0; and  $\tau \notin S_{\tau,\hat{p}}^\infty$  given that  $\tau$  is not a Nash distribution. Since  $p' \geq \hat{p}$  implies  $S_{\tau,p'}^\infty \subset S_{\tau,\hat{p}}^\infty$ , the result follows.

(ii) Consider an open neighborhood  $N$  of  $\tau^*$  in  $S_{\tau^*,p}^\infty$ . By continuity of  $\phi$  and definition of  $S_{\tau^*,p}^\infty$ , we have that  $S_{\tau^*,p}^\infty = \phi(S_{\tau^*,p}^\infty) \cap T_{\tau^*,p}$ . This means that for every  $\tau \in N$ , there exists a non empty set  $B_{\tau^*,p}(\tau)$  of distributions  $\tau' \in S_{\tau^*,p}^\infty$  such that  $\phi(p\tau^* + (1-p')\tau') = \tau$ . Let  $B_{\tau^*,p} = \cup_{\tau \in N} B_{\tau^*,p}(\tau)$ . Note that  $B_{\tau^*,p} \subset S_{\tau^*,p}^\infty$  and  $\tau^* \in B_{\tau^*,p}$  (since  $\tau^* \in B_{\tau^*,p}(\tau^*)$ ). Let  $T_{\tau^*,p}(N) = \{p\tau^* + (1-p)\tau', \tau' \in N\}$ . Note that  $T_{\tau^*,p}(B_{\tau^*,p}) \subseteq \text{Int}.T_{\tau^*,p}$ ; hence, for any  $\tau$  in  $N$  (provided  $N$  is chosen small enough),  $T_{\tau^*,p}(B_{\tau^*,p}) \subseteq \text{Int}.T_{\tau,p}$ . Therefore,  $T_{\tau^*,p}(B_{\tau^*,p}) = T_{\tau,p}(B_{\tau,p})$  so that, by continuity of  $\phi$ ,  $B_{\tau,p} \subseteq S_{\tau^*,p}^\infty$  so that  $N \subseteq S_{\tau^*,p}^\infty$ , as required. ■

An informal interpretation of Point (i) is that under a mild continuity restriction (which would typically be satisfied, for example, in smooth settings (see Section 3.2 below)), a  $p$ -stable equilibrium is a locally isolated  $p$ -consensus distribution, *i.e.* there is no  $p'$ -consensus distribution in its vicinity with a degree of consensus  $p'$  larger than  $p$ . Informally again, Point (ii) says the converse statement also holds provided the best-response map  $\phi$  is continuous: in the vicinity of a  $p$ -stable equilibrium, there is always at least one  $p'$ -consensus distribution with a degree of consensus  $p'$  smaller than  $p$  and arbitrarily close to  $p$  (provided that the interior of  $S_{\tau^*,p}^\infty$  is not empty).

However, note that these informal interpretations are offered to aid understanding of Proposition 3; the analogous formal statements cannot, in general, be proved. Indeed, the (very simple) game developed in Section 3.1 is a counterexample: there is a  $p$ -stable equilibrium such that every distribution  $\tau$  in its neighborhood is a  $p_\tau$ -consensus distribution for some value  $p_\tau > p$  ( $p_\tau$  depends on  $\tau$ ); there is another  $p$ -stable equilibrium with no  $p_\tau$ -consensus distribution in its neighborhood satisfying  $p_\tau \geq p$ . The point is that  $p_\tau$  tends to  $p$  when  $\tau$  tends to the equilibrium.

Lastly, the requirement that  $S_{\tau^*,p}^\infty$  has a non-empty interior in  $T$  is a restriction that is key to Proposition 3 and may not always be satisfied when the action set has at least two dimensions. To see this point, consider the simple case of a game where an action is a vector in  $\mathfrak{R}^K$  for some  $K \geq 2$  and the utility depends only on the individual action and the average action (*i.e.* only the first moment of the distribution of actions matter for the strategic interaction). The notation  $S_{\tau,p}^\infty$  and  $\tau$  are easily redefined as sets and elements in  $\mathfrak{R}^K$ . In such a

game, it may be the case that  $S_{\tau^*,p}^\infty$  has a dimension strictly less than  $K$  (hence, an empty interior). For  $\tau$  close to  $\tau^*$ , the set  $S_{\tau,p}^\infty$  may have a dimension strictly less than  $K$  as well (it may be close to  $S_{\tau^*,p}^\infty$  by a continuity argument). Because of the low dimension of  $S_{\tau,p}^\infty$ , it is fully possible that  $\tau$  is close to  $\tau^*$ ,  $S_{\tau,p}^\infty$  is close to  $S_{\tau^*,p}^\infty$  and yet  $\tau \notin S_{\tau,p}^\infty$ . Consequently, there may be no  $\tau$  in the vicinity of  $\tau^*$  that are  $p$ -consensus.

### 3 Illustrative examples

In this section, we demonstrate how the two concepts developed here apply in a number of illustrative examples (one dimensional linear and smooth settings) and examine the link between the two concepts and a number of other related concepts studied in finite games.

#### 3.1 Games with linear one-dimensional best responses

As  $p$ -consensus is a refinement of rationalizability, we might expect a link between the set of  $p$ -consensus distributions and subjective correlated equilibria. To obtain a clear link with simple correlated equilibria, we work in a simpler framework where payoffs depend on own action and the average action.

A game with a continuum of players  $i \in [0, 1]$ , actions  $a \in [-1, 1]$  and every player has the same BR map as follows:

$$\begin{aligned} a_i &= \beta \bar{a} \text{ for } \bar{a} \in [-1/|\beta|, 1/|\beta|], \\ a_i &= -\frac{\beta}{|\beta|} \text{ for } \bar{a} < -1/|\beta|, \\ a_i &= \frac{\beta}{|\beta|} \text{ for } \bar{a} > 1/|\beta|, \end{aligned}$$

where  $\bar{a}$  denotes the average action and  $\beta$  is a real parameter ( $\beta > 0$  is the case with strategic complements and  $\beta < 0$  the case with strategic substitutes).  $\frac{\beta}{|\beta|}$  is either  $-1$  or  $1$  depending on the sign of  $\beta$ .

In what follows, with some abuse of notation and ease of exposition, we use  $a$  to denote the Dirac measure on  $a$  for any action  $a \in [-1, 1]$  and we call an average action  $\bar{a}$  a rationalizable or a  $p$ -consensus outcome when it corresponds to a rationalizable or a  $p$ -consensus distribution of individual actions.

**p-stability.** When  $|\beta| < 1$ , the unique rationalizable outcome is the (unique) Nash equilibrium  $0$  (the equilibrium is  $0$ -stable). When  $|\beta| > 1$ , the result differs according to the sign of  $\beta$ .

When  $\beta > 1$ , there are three Nash equilibria:  $-1, 0, 1$ . Then, by computation, it follows that (i) the equilibrium  $0$  is  $p$ -stable for  $p = 1 - \frac{1}{\beta}$ , (ii) the equilibria  $-1$  and  $1$  are  $p$ -stable for  $p = \frac{1}{2} \left(1 - \frac{1}{\beta}\right)$ .

When  $\beta < -1$ , there is still exactly one Nash equilibrium:  $0$ . It is  $p$ -stable for  $p = 1 - \frac{1}{|\beta|}$ .

Note that for the equilibrium 0, the upper bound  $K_{\tau^*}$  introduced in Definition 3 (to check admissibility of an equilibrium) is  $|\beta|$ . For the corner equilibria  $-1$  and  $1$ , this upper bound is 0. Hence, every equilibrium in this game is admissible and it is  $p$ -stable for some  $p < 1$  (by Proposition 2).

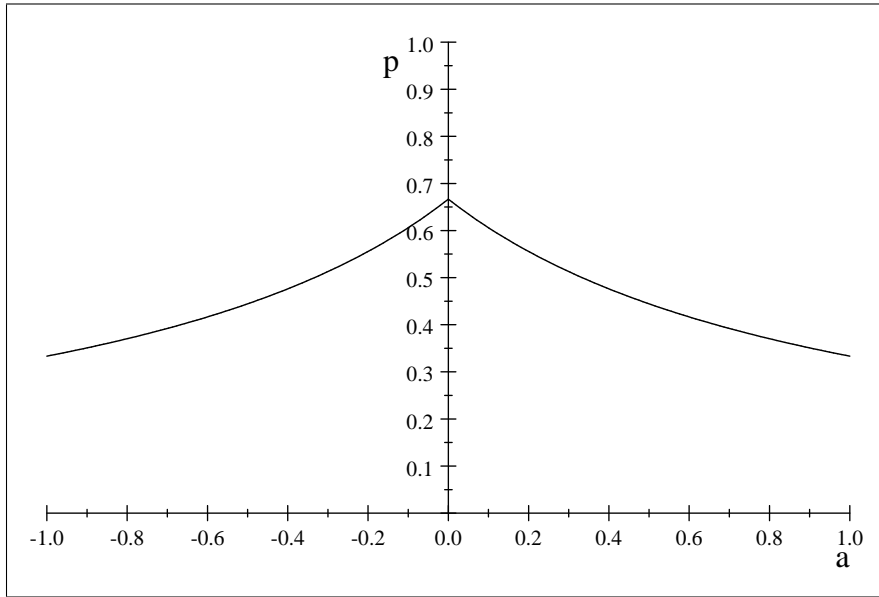
**$p$ -consensus.** We compute the  $p$ -consensus outcomes that differ from Nash equilibria. If  $|\beta| < 1$ , no outcome different from a Nash equilibrium is a  $p$ -consensus outcome, and we restrict attention to the case  $|\beta| > 1$  in what follows. The results in the cases with strategic complements and strategic substitutes are slightly different.

**Proposition 4a (strategic complements).** For  $\beta > 1$ , every  $\bar{a}$  that differ from any Nash equilibrium ( $\bar{a} \notin \{-1, 0, 1\}$ ) is a  $p$ -consensus outcome for

$$p = \frac{1}{1 + |\bar{a}|} \left(1 - \frac{1}{\beta}\right) \in \left[\frac{1}{2} \left(1 - \frac{1}{\beta}\right), 1 - \frac{1}{\beta}\right].$$

**Proof.** See appendix. ■

Observe that every outcome is a  $p$ -consensus outcome for a value  $p$  between the degree  $\frac{1}{2} \left(1 - \frac{1}{\beta}\right)$  of  $p$ -stability of the corner equilibria  $\pm 1$  and the degree  $\left(1 - \frac{1}{\beta}\right)$  of  $p$ -stability of the interior equilibrium 0. This result illustrates Proposition 3.



The logic of the proof is analogous to the well-known argument characterizing the rationalizable outcomes (as the outcomes between the lowest and the highest fixed points of the best-response map): The  $p$ -BR map is increasing and the bounds of the set  $S_{\bar{a}, p}^\infty$  are the lowest and the highest fixed points of the  $p$ -BR

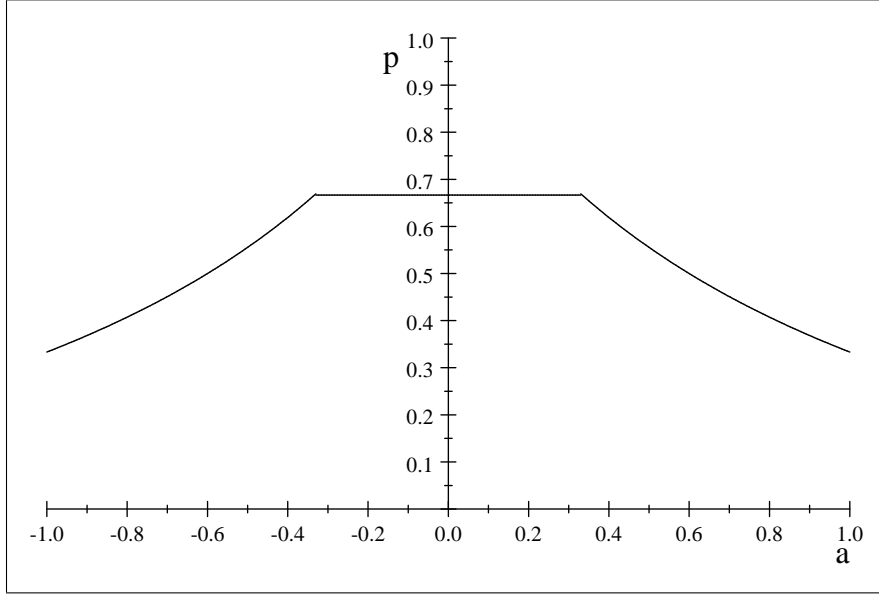
map; since the  $p$ -BR map has either one fixed point or 3 fixed points (including  $-1$  and  $1$ ), the set  $S_{\bar{a},p}^\infty$  is either reduced to a single point (either  $-1$  or  $1$  or  $\frac{\beta p \bar{a}}{1-\beta(1-p)}$ , and  $\bar{a} \notin S_{\bar{a},p}^\infty$ ) or equal to  $[-1, 1]$  (and  $\bar{a} \in S_{\bar{a},p}^\infty$ ).

**Proposition 4b (strategic substitutes).** For  $\beta < -1$ , every  $\bar{a}$  that differ from the Nash equilibrium ( $\bar{a} \neq 0$ ) is a  $p$ -consensus outcome for

$$p = \min \left( \frac{1}{1+|\bar{a}|} \left( 1 + \frac{|\bar{a}|}{\beta} \right), 1 + \frac{1}{\beta} \right) \in \left[ \frac{1}{2} \left( 1 + \frac{1}{\beta} \right), 1 + \frac{1}{\beta} \right].$$

**Proof.** See appendix. ■

Observe that every outcome is a  $p$ -consensus outcome for a value  $p$  lower than the degree  $\left(1 - \frac{1}{\beta}\right)$  of  $p$ -stability of the Nash equilibrium.



$p$ -consensus in the case  $\beta < -1$  (example with  $\beta = -3$ )

Again, the logic of the proof is analogous to the argument characterizing the rationalizable outcomes (as the outcomes between the 2 points of the largest 2-cycle, *i.e.*, the lowest and the highest fixed points of the 2nd iterate of the best-response map): The second iterate of the  $p$ -BR map is increasing and the bounds of the set  $S_{\bar{a},p}^\infty$  are the lowest and the highest fixed points of this second iterate; since the second iterate map has either one fixed point or 3 fixed points, the set  $S_{\bar{a},p}^\infty$  is either reduced to a single point (and  $\bar{a} \notin S_{\bar{a},p}^\infty$ ) or equal to an interval. The difference with the case with strategic complements is that the bounds of this interval (the fixed points of the 2nd iterate) are not necessarily  $-1$  and  $1$ . Then, the condition  $\bar{a} \in S_{\bar{a},p}^\infty$  is not the same as in the case with strategic complements.

**Correlated Equilibrium.** Consider a  $p$ -consensus outcome  $\bar{a}$ . Clearly, for any  $p' > p$ , there cannot exist a correlated equilibrium such that (i)  $\bar{a}$  is

one of the possible outcomes and (ii) at every state, conditionally to his/her information, everyone assigns a probability at least  $p'$  to  $\bar{a}$ .<sup>4</sup> This follows from the fact that every outcome of this correlated equilibrium should belong to  $S_{\bar{a},p'}^\infty$ , which is impossible since  $\bar{a} \notin S_{\bar{a},p'}^\infty$ . We show the converse statement: we exhibit a simple correlated equilibrium (2 outcomes only) such that (i)  $\bar{a}$  is one of the possible outcomes and (ii) at every state, everyone assigns a probability at least  $p$  to  $\bar{a}$ .

The case  $\beta > 0$  is the simple case, while  $\beta < 0$  is less convenient to solve.

Consider first  $\beta > 0$  and an outcome  $\bar{a} > 0$ .  $\bar{a}$  is a  $p$ -consensus for the value of  $p$  defined in Proposition 4a. We define the correlated equilibrium as follows. The 2 outcomes are  $\bar{a}$  (with probability  $\pi$ , defined below) and  $-1$  (with probability  $1 - \pi$ ). Individual belief is either the "informed" belief (probability 1 on  $\bar{a}$ ) or the "uninformed" belief " $\bar{a}$  with probability  $p$  and  $-1$  with probability  $1 - p$ ". When the outcome is  $-1$ , every agent holds the "uninformed" belief. When the outcome is  $\bar{a}$ , a proportion  $\alpha$  of agents (defined below) holds the "informed" belief and the remaining proportion holds the "uninformed" belief.

We now check that this distribution of actions is a correlated equilibrium. The optimal action of an "uninformed" is  $-1$  since (given the value of  $p$ )

$$-1 \geq \beta(p\bar{a} + (1-p)(-1)).$$

Hence, when the outcome is  $-1$ , the optimal action of every agent is  $-1$ , and the distribution of belief is consistent with the outcome.

We distinguish between 2 subcases.

In the subcase  $\bar{a} \leq 1/\beta$ , let

$$\pi = \frac{1}{\bar{a} + 1} \text{ and } \alpha = \frac{\bar{a} + 1}{\beta\bar{a} + 1} \in (0, 1).$$

The optimal action of an "informed" is  $\beta\bar{a}$ . Hence, when the outcome is  $\bar{a}$ , the optimal action of every agent is either  $-1$  or  $\beta\bar{a}$ , the resulting average action is

$$\alpha\beta\bar{a} + (1 - \alpha)(-1) = \bar{a},$$

and the distribution of belief is consistent with the outcome. Lastly, the "uninformed" belief is consistent with the prior distribution since Bayes rule writes

$$p = \frac{\pi(1-\alpha)}{\pi(1-\alpha) + 1 - \pi}.$$

In the subcase  $1 > \bar{a} \geq 1/\beta$ , let

$$\pi = \frac{\frac{1}{1+\bar{a}} \left(1 - \frac{1}{\beta}\right)}{\frac{1}{2} \left(1 - \frac{1}{\beta}\right) + \frac{1-\bar{a}}{2}} \in (0, 1) \text{ and } \alpha = \frac{\bar{a} + 1}{2} \in (0, 1).$$

The optimal action of an "informed" is 1. Hence, when the outcome is  $\bar{a}$ , the optimal action of every agent is either  $-1$  or 1, the resulting average action is

$$\alpha + (1 - \alpha)(-1) = \bar{a},$$

---

<sup>4</sup>Even a weak form of correlated equilibrium, like subjective correlated equilibrium cannot satisfy these properties.

and the distribution of belief is consistent with the outcome. Lastly, the "uninformed" belief is consistent with the prior distribution since Bayes rule writes  $p = \frac{\pi(1-\alpha)}{\pi(1-\alpha)+1-\pi}$ .

The correlated equilibrium with  $\bar{a} < 0$  is defined analogously (with +1 for the second outcome).

We now consider  $\beta < 0$  and an outcome  $\bar{a} > 0$ .  $\bar{a}$  is a  $p$ -consensus for the value of  $p$  defined in Proposition 4b. We distinguish between the 2 subcases  $\bar{a} \leq -1/\beta$  and  $\bar{a} > -1/\beta$ .

In the first subcase  $\bar{a} \leq -1/\beta$ , for any small enough  $\varepsilon > 0$ , we define the correlated equilibrium as follows. The 2 outcomes are  $\bar{a}$  and some other outcome  $\bar{a}'$ . There are 2 types of agents: Either an agent holds the "uninformed" belief ( $\bar{a}$  with probability  $p - \varepsilon$  and  $\bar{a}'$  with probability  $1 - p + \varepsilon$ ) and plays an action  $\hat{a}_U$  (defined below), or an agent holds the "almost informed" belief ( $\bar{a}$  with probability  $q > p$  and  $\bar{a}'$  with probability  $1 - q$ ) and plays the action  $\hat{a}_I$  (defined below). When the outcome is  $\bar{a}$ , a proportion  $\alpha$  of agents (defined below) holds the "uninformed" belief and the remaining proportion holds the "almost informed" belief. When the outcome is  $\bar{a}'$ , the proportion of agents holding the "uninformed" belief is  $\alpha'$  (defined below).

We now write the conditions ensuring that this distribution of actions and beliefs is a correlated equilibrium. The optimal action  $\hat{a}_U$  of an "uninformed" satisfies

$$\hat{a}_U = \beta((p - \varepsilon)\bar{a} + (1 - p + \varepsilon)\bar{a}') \in (-1, 1),$$

and the optimal action  $\hat{a}_I$  of an "almost informed" satisfies

$$\hat{a}_I = \beta(q\bar{a} + (1 - q)\bar{a}') \in (-1, 1).$$

When the outcome is  $\bar{a}$ , given the distribution of beliefs in the population, the equilibrium condition is

$$\bar{a} = \alpha\hat{a}_I + (1 - \alpha)\hat{a}_U,$$

and, when the outcome is  $\bar{a}'$ , the equilibrium condition is

$$\bar{a}' = \alpha'\hat{a}_I + (1 - \alpha')\hat{a}_U.$$

An equilibrium is obtained for  $(\alpha, \alpha', q, \bar{a}', \hat{a}_U, \hat{a}_I)$  satisfying

$$\begin{aligned} \bar{a}' &\in \left( \beta\bar{a}, \frac{(1 - \varepsilon)\beta}{1 - \beta\varepsilon}\bar{a} \right) \subseteq (-1, 0), \\ q &\in \left( \left(1 - \frac{1}{\beta}\right) \frac{\bar{a}'}{\bar{a}' - \bar{a}}, 1 \right) \end{aligned}$$

and

$$\begin{aligned} \hat{a}_U &= \beta((p - \varepsilon)\bar{a} + (1 - p + \varepsilon)\bar{a}') \in (\bar{a}, 1), \\ \hat{a}_I &= \beta(q\bar{a} + (1 - q)\bar{a}') \in (-1, \bar{a}'), \\ \alpha &= \frac{\bar{a} - \hat{a}_U}{\hat{a}_I - \hat{a}_U} \in (0, 1), \\ \alpha' &= \frac{\bar{a}' - \hat{a}_U}{\hat{a}_I - \hat{a}_U} \in (0, 1). \end{aligned}$$

(note that there exist parameters values consistent with these conditions since  $-1 \leq \beta\bar{a} < \frac{(1-\varepsilon)\beta}{1-\beta\varepsilon}\bar{a} < 0$  and  $(1 - \frac{1}{\beta}) \frac{\bar{a}'}{\bar{a}' - \bar{a}} < 1$ ).

**Proof.** Given that  $p = 1 + \frac{1}{\beta}$ ,  $\beta < -1$  and  $-1/\beta \geq \bar{a} > 0$ , computations show  $\hat{a}_U < 1$  for  $\varepsilon$  small enough (note that  $\beta\bar{a} + \bar{a} - \bar{a}' < -\bar{a}' < 1$ ) and  $\hat{a}_U > \bar{a}$  follows from  $\bar{a}' < \frac{(1-\varepsilon)\beta}{1-\beta\varepsilon}\bar{a}$ . Hence,  $\hat{a}_U$  is the optimal action for an "uninformed" agent. Computations show  $\hat{a}_I > -1$  (since  $q\bar{a} + (1-q)\bar{a}' < \bar{a} \leq -\frac{1}{\beta}$ ) and  $\hat{a}_I < \bar{a}'$  follows from the lower bound on  $q$ . Hence,  $\hat{a}_I$  is the optimal action for an "informed" agent. Lastly,  $\alpha$  and  $\alpha'$  satisfy the equilibrium conditions. The conditions  $\hat{a}_U \in (\bar{a}, 1)$  and  $\hat{a}_I \in (-1, \bar{a}')$  imply  $\alpha, \alpha' \in (0, 1)$ . ■

A heterogenous prior distribution can be defined.

In the second subcase  $\bar{a} > -1/\beta$ ,  $p = \frac{1}{1+\bar{a}} \left(1 + \frac{\bar{a}}{\beta}\right)$ , for any small enough  $\varepsilon > 0$ , we define the correlated equilibrium as follows. The 2 outcomes are  $\bar{a}$  and some other outcome  $\bar{a}'$ . There are 2 types of agents: Either an agent holds the "uninformed" belief ( $\bar{a}$  with probability  $p - \varepsilon$  and  $\bar{a}'$  with probability  $1 - p + \varepsilon$ ) and plays an action  $\hat{a}_U$  (defined below), or an agent holds the "almost informed" belief ( $\bar{a}$  with probability  $q > p$  and  $\bar{a}'$  with probability  $1 - q$ ) and plays the action  $-1$ . When the outcome is  $\bar{a}$ , a proportion  $\alpha$  of agents (defined below) holds the "uninformed" belief and the remaining proportion holds the "almost informed" belief. When the outcome is  $\bar{a}'$ , the proportion of agents holding the "uninformed" belief is  $\alpha'$  (defined below).

We now write the conditions ensuring that this distribution of actions and beliefs is a correlated equilibrium. The optimal action  $\hat{a}_U$  of an "uninformed" satisfies

$$\hat{a}_U = \beta((p - \varepsilon)\bar{a} + (1 - p + \varepsilon)\bar{a}') \in (-1, 1),$$

and the optimal action of an "almost informed" is  $-1$  if

$$-1 \geq \beta(q\bar{a} + (1 - q)\bar{a}').$$

When the outcome is  $\bar{a}$ , given the distribution of beliefs in the population, the equilibrium condition is

$$\bar{a} = \alpha(-1) + (1 - \alpha)\hat{a}_U,$$

and, when the outcome is  $\bar{a}'$ , the equilibrium condition is

$$\bar{a}' = \alpha'(-1) + (1 - \alpha')\hat{a}_U.$$

An equilibrium is obtained for  $(\alpha, \alpha', q, \bar{a}', \hat{a}_U)$  satisfying

$$\bar{a}' \in \left(-1, \frac{1 - \beta(p - \varepsilon)}{\beta(1 - p + \varepsilon)}\bar{a}\right) \subseteq (-1, 0),$$

$$q \in \left(-\frac{\frac{1}{\beta} + \bar{a}'}{\bar{a} - \bar{a}'}, 1\right)$$

and

$$\begin{aligned}\hat{a}_U &= \beta((p - \varepsilon)\bar{a} + (1 - p + \varepsilon)\bar{a}') \in (\bar{a}, 1) \\ \alpha &= \frac{\hat{a}_U - \bar{a}}{\hat{a}_U + 1} \in (0, 1), \\ \alpha' &= \frac{\hat{a}_U - \bar{a}'}{\hat{a}_U + 1} \in (0, 1).\end{aligned}$$

(note that there exist parameters values consistent with these conditions since  $-1 < \frac{1 - \beta(p - \varepsilon)\bar{a}}{\beta(1 - p + \varepsilon)} < 0$  and  $-\frac{1}{\beta} < \frac{\bar{a}'}{\bar{a} - \bar{a}'} < 1$ ).

**Proof.** Given that  $p = \frac{1}{1 + \bar{a}} \left(1 + \frac{\bar{a}}{\beta}\right)$ ,  $\beta < -1$  and  $\bar{a} > -1/\beta$ , computations show  $\hat{a}_U < 1$  for  $\varepsilon$  small enough (note that  $\bar{a}(\beta - 1)\bar{a}' < \bar{a}(1 - \beta) < (1 + \bar{a}) - (\beta + \bar{a})\bar{a}$ ) and  $\hat{a}_U > \bar{a}$  follows from  $\bar{a}' < \frac{1 - \beta(p - \varepsilon)}{\beta(1 - p + \varepsilon)}\bar{a}$ . Hence,  $\hat{a}_U$  is the optimal action for an "uninformed" agent. The condition on  $q$  implies that the optimal action of the "informed" is  $-1$ . Lastly,  $\alpha$  and  $\alpha'$  satisfy the equilibrium conditions. The conditions  $\hat{a}_U > \bar{a}$  and  $\bar{a}' < 0$  imply  $\alpha, \alpha' \in (0, 1)$ . ■

The last case ( $\beta < 0$  and  $\bar{a} \leq 0$ ) is analogous to the above case  $\beta < 0$  and  $\bar{a} \geq 0$ , using  $+1$  instead of  $-1$  as a 2nd outcome when needed (to check this point, one just has to multiply every equilibrium condition by  $-1$ ).

### 3.2 p-stability in the smooth one-dimensional case

To get an intuitive feel for the notion of stability being studied in this paper, we extend the piecewise linear setting of the previous subsection to a non linear setting.

Consider a simple, smooth model of strategic interaction where there is a continuum of agents each whom chooses an action  $a \in A$  (a compact set in  $R$ ) to maximize  $u(a, \bar{a})$  ( $C^2$ , with  $u''_{aa} < 0$ ) where  $\bar{a}$  is the average action. Without loss of generality,  $A = [-1, 1]$ . Suppose, there is a (not necessarily) unique Nash that is interior and is normalized to 0 so that  $u'_a(0, 0) = 0$ . Denote  $BR(\bar{a})$  the (unique) best response to  $\bar{a}$  (characterized by  $u'_a(BR(\bar{a}), \bar{a}) = 0$ ). We assume that the  $BR$  map is not vertical at equilibrium ( $BR'(0) < +\infty$ ).

We are now in a position to state the following result:

**Proposition 5.** There is  $\hat{p} < 1$  such that the equilibrium is  $\hat{p}$ -stable. If  $\sup_{\bar{a} \in [-1, 1]} |BR'(\bar{a})| < 1$ , then  $\hat{p} = 0$ . Otherwise, we have:

$$1 - \frac{1}{|BR'(0)|} \leq \hat{p} \leq 1 - \frac{1}{1 + (M - 1)m}, \quad (1)$$

where

$$\begin{aligned}m &= \sup_{a, \bar{a} \in [-1, 1]} \frac{u''_{a\bar{a}}(a, \bar{a})}{u''_{a\bar{a}}(a, 0)} \geq 1, \\ M &= \sup_{a, \bar{a} \in [-1, 1]} \left| \frac{u''_{a\bar{a}}(a, \bar{a})}{u''_{aa}(a, \bar{a})} \right| \geq 1.\end{aligned}$$



For any  $p < \hat{p}$ , there exists a neighborhood of 0 such that every action in this neighborhood is the average action of a  $p$ -consensus distribution. For any  $p > \hat{p}$ , there exists a neighborhood of 0 such that no action in this neighborhood is the average action of a  $p$ -consensus distribution.

**Proof.** See appendix. ■

Notice that by implicit functions theorem,  $|BR'(\bar{a})| \leq M$  for any  $\bar{a}$ . To relate  $\hat{p}$  with exogenous variables, rewrite the left inequality (1) using  $BR'(0) = -u''_{a\bar{a}}(0,0)/u''_{aa}(0,0)$  (by implicit functions theorem, again). The linear case developed in the previous subsection corresponds to the case with a quadratic utility:  $m = 1$ ,  $BR'$  is constant (equal to  $M$ ) and  $\hat{p} = 1 - 1/M$ . As in the linear case, our stability concept gives a motivation for looking at the slope of the best response map as a stability index.

A special case of the model studied so far is the Muth model with a large number of farmers who have to commit to an output level before selling their products in a competitive market in Guesnerie (1992). Farmer  $i$  maximizes  $\pi q - \frac{q^2}{2C^i}$  ( $\pi$  is the output price). Aggregate supply in this market is given by  $S(\pi) = C\pi$  where  $C = \int C^i di$ . Aggregate demand in this market is:

$$D(\pi) = \begin{cases} A - B\pi & \text{if } \pi \leq \frac{A}{B} \\ 0, & \text{otherwise} \end{cases}$$

Let  $\pi^*$  be the competitive equilibrium price. Guesnerie (1992) shows that when the slope of the best response map  $B/C < 1$ ,  $\pi^*$  is the unique rationalizable outcome.

Applying Proposition 5 immediately yields that the equilibrium in Guesnerie's model is  $\hat{p}$ -stable for  $\hat{p} = \max\{1 - C/B, 0\}$ . Thus,  $p$ -stability describes more precisely the degree of stability of the equilibrium when it is not the unique rationalizable outcome.

The remainder of the section is devoted to the proof of Proposition 5. The proof shows that  $p$ -stability relies on the best response map  $BR_{0,p}(\bar{a})$  (best response to beliefs "probability  $p$  on 0, probability  $(1 - p)$  on  $\bar{a}$ "). When  $p$  is close to one, the slope  $BR'_{0,p}(\bar{a})$  is small enough (whatever  $\bar{a}$  is). The map  $BR_{0,p}$  is then globally contracting and  $p$ -stability obtains. Intuitively, when  $p$  is close to one, the best response is not very sensible to the value  $\bar{a}$  and the best response cannot deviate very much from the equilibrium value 0. This is the condition needed to get  $p$ -stability.

### 3.3 Basin of attraction and $p$ -stability

In this subsection, we show, in an example with three equilibria, that the  $p$ -stability of a corner equilibrium is linked to the slope of the aggregate best response at that equilibrium but not to the size of the basin of attraction.

We consider a game with a continuum of individuals of mass one who must choose an action  $a \in [-1, 1]$  to maximize the payoffs

$$-\frac{a^2}{2} + aP(\bar{a})$$

where  $\bar{a}$  denotes the average action and  $P(\cdot)$  is a third degree polynomial specified below. At an interior best response, the first order condition is  $P(\bar{a}) - a = 0$ .

Assume that  $P(\bar{a}) = -A(\bar{a} - 1)(\bar{a} + 1)(\bar{a} - \alpha)$ , where  $A > 0$  and  $\alpha \in (-1, 1)$ . Then, there are three equilibria (fix points):  $\bar{a} = 1$ ,  $\bar{a} = -1$ ,  $\bar{a} = \alpha$ . By computation, we can check that

$$\left. \frac{\partial P(\bar{a})}{\partial \bar{a}} \right|_{\bar{a}=1} = 1 - 2A(1 - \alpha), \quad \left. \frac{\partial P(\bar{a})}{\partial \bar{a}} \right|_{\bar{a}=-1} = 1 - 2A(1 + \alpha), \quad \left. \frac{\partial P(\bar{a})}{\partial \bar{a}} \right|_{\bar{a}=\alpha} = 1 + A(1 - \alpha^2).$$

Note that  $\left. \frac{\partial P(\bar{a})}{\partial \bar{a}} \right|_{\bar{a}=\alpha} > 1$  while  $\left. \frac{\partial P(\bar{a})}{\partial \bar{a}} \right|_{\bar{a}=1}$  and  $\left. \frac{\partial P(\bar{a})}{\partial \bar{a}} \right|_{\bar{a}=-1}$  are both less than one. To ensure that we are in the case of strategic complements we restrict the parameters so that both  $\left. \frac{\partial P(\bar{a})}{\partial \bar{a}} \right|_{\bar{a}=1} > 0$  and  $\left. \frac{\partial P(\bar{a})}{\partial \bar{a}} \right|_{\bar{a}=-1} > 0$ . Therefore,  $\bar{a} = 1$  (respectively,  $\bar{a} = -1$ ) is stable in the best-response dynamics on its basin of attraction  $(\alpha, 1]$  (respectively,  $[-1, \alpha)$ ). Furthermore, it follows that the whole action set is rationalizable so, in particular, none of the three equilibria is 0-stable.

Next, we show that by choosing different value of  $A$  we can choose different values of  $\alpha$  consistent with the  $p$ -stability of  $\bar{a} = 1$ . By computation, the  $p$ -best response map is

$$pP(\bar{a}) + (1 - p)P(\bar{a}') = pP(\bar{a}) - A[(1 - p)(\bar{a}' - 1)(\bar{a}' + 1)(\bar{a}' - \alpha)] + (1 - p)\bar{a}'$$

Observe that  $p$ -stability requires the above best-response map to be convergent and we need a unique fix point of the preceding map. So we look for value of  $p$  for which the preceding  $p$ -best response has exactly one root. So the requiring  $p$ -stability of either one of the two equilibria for all  $p > \bar{p}$  requires us to calculate the value of  $p$  for which the preceding  $p$ -best response has exactly two roots one of which is the equilibrium and the other one is the double root. So this implies

$$pP(\bar{a}) + (1 - p)P(\bar{a}') = -A(1 - p)(\bar{a}' - \bar{a})(\bar{a}' - \delta)^2$$

where  $\delta$  is the unknown double root.

Consider the  $p$ -stability of  $\bar{x} = 1$ . Identifying the coefficients (the coefficients of degree 3 are identical)

$$\begin{aligned} (1 - p)\alpha A &= (1 - p)A(1 + 2\delta) \\ (1 - p)(1 + A) &= 1 - (1 - p)A(2 + \delta)\delta \\ p - (1 - p)\alpha A &= (1 - p)A\delta^2 \end{aligned}$$

The first one requires  $\delta = \frac{\alpha - 1}{2}$  while the second one requires

$$(1 - p)A(1 + \delta)^2 = p$$

which is  $\delta = \sqrt{\frac{p}{(1 - p)A}} - 1$  (the second one follows from the others). Hence,  $p$  is

such that

$$\frac{\alpha - 1}{2} = \sqrt{\frac{p}{(1-p)A}} - 1$$

$$\frac{A(\alpha + 1)^2}{4} = \frac{p}{(1-p)}$$

Observe that we can rewrite the preceding expression as

$$\frac{\left(1 - \frac{\partial P(\bar{a})}{\partial \bar{a}} \Big|_{\bar{a}=1}\right) (\alpha + 1)^2}{8(1 - \alpha)} = \frac{p}{(1-p)}$$

It follows that by choosing different value of  $A$  we can choose different values of  $\alpha$  consistent with the  $p$ -stability of  $\bar{a} = 1$ .

Therefore, there is no link between  $p$ -stability of  $\bar{a} = 1$  and the size of its basin of attraction.

## 4 $p$ -consensus and $p$ -stability in finite games

In this subsection, we relate the two notions of  $p$ -consensus and  $p$ -stability to a number of related concepts developed in games with a finite number of players and (pure) actions.

It is evident that an equilibrium strategy profile with a positive measure of players with more than 1 best response to it cannot be  $p$ -stable. Consistent with common knowledge of the equilibrium outcome, a player can choose a strategy different from the equilibrium if the player expects other players' to do so: the belief that the equilibrium outcome is common knowledge does not imply that a player chooses the equilibrium strategy as a best response to the equilibrium outcome. Hence, in a finite game, no mixed strategy equilibrium or a pure strategy equilibrium in weakly dominated strategies can be  $p$ -stable. The same point applies to games with a finite number of players but a continuum of actions where there are several best-responses to an equilibrium outcome, as in a second-price auction where truth-telling (an equilibrium in weakly dominated strategies) isn't  $p$ -stable. In many mechanisms (tournaments, voting games) used to implement socially optimal outcomes as a weakly dominant strategy, whether a player is a winner/loser doesn't depend on the action of the player., In such cases, equilibrium outcomes are "structurally" supported by several best responses: as the equilibrium actions profile implies a profile of winner/losers, which defines payoffs to any individual player, a player's payoff as a winner/loser does not depend on profile of winner/losers over other players.

Next, we point out that the approximate common of beliefs (and the corresponding concept of belief potential) is quite different from the common knowledge of  $p$ -beliefs (which is needed for the definition of  $p$ -consensus and  $p$ -stability) because approximate common knowledge of a Nash equilibrium profile of actions is consistent with the fact that the Nash equilibrium profile of

actions is not  $p$ -believed anywhere (while this is required for the definition of  $p$ -consensus and  $p$ -stability). The following example illustrates the preceding point:

**Example.** There are two players and each player has three actions in the following symmetric game (call the actions:  $Up(U)$ ,  $Middle(M)$ ,  $Bottom(B)$ ):

$$\begin{bmatrix} & U & M & B \\ U & 1, 1 & 1, 0 & 0, -1 \\ M & 0, 1 & -1, -1 & 1, 0 \\ B & -1, 0 & 0, 1 & 1, 1 \end{bmatrix}$$

We show: (a)  $(U, U)$  is a  $\frac{1}{3}$ -stable Nash equilibrium, (b) the existence of a correlated equilibrium with belief potential (see Morris, Rob and Shin (1995) for a formal definition) of  $\frac{1}{3}$  and  $(U, U)$  is played with probability strictly greater than  $\frac{1}{3}$  but other actions are played as well. Ex-ante, it is common knowledge that  $(U, U)$  is played with probability strictly greater than  $\frac{1}{3}$ , but the interim information provided on the other player's action is such that  $(U, U)$  is not necessarily played (at the interim stage, it is not always true that  $(U, U)$  is  $\frac{1}{3}$ -believed).for  $p = \frac{1}{3}$ -stable.

By computation, observe that  $(U, U)$  is a  $p$ -stable Nash equilibrium for  $p = 1/3$ -stable:  $S_0^{\frac{1}{3}} = \{(U, U), (M, M), (B, B)\}$ ,  $S_1^{\frac{1}{3}} = \{(U, U), (M, M)\}$ ,  $S_2^{\frac{1}{3}} = \{(U, U)\}$ . For  $B$  to be a best response and not to be in  $S_1^p$ , we need that for each player  $B$  is a best response to beliefs putting some probability on  $(B, b)$ . For  $M$  to be in  $S_1^{\frac{1}{3}}$  and not  $S_2^{\frac{1}{3}}$ , it is required that  $M$  is a best response to beliefs putting some probability on  $B$  (so that  $U$  cannot be a best response to  $B$ ) while  $U$  strictly dominates  $M$  for beliefs putting probability 1 on  $\{U, M\}$ . Hence,  $1/3$ -stability obtains as soon as  $B$  is not a best-response to any  $1/3$ -beliefs on  $U$  (which is guaranteed by the fact that  $B$  is not a best response to  $U$ ). Next, we construct a correlated equilibrium with belief potential  $1/3$ . We want a correlated equilibrium with belief potential  $p = \frac{1}{3}$  where a player plays different actions (not only  $U$ ). At the equilibrium, the set of the actions played at a correlated equilibrium is common knowledge; hence this set cannot be restricted to 2 actions (otherwise, given the belief potential,  $U$  is the unique best response at every state: a contradiction). Hence, the set of actions played by each player at the correlated equilibrium is  $\{U, M, B\}$ . The construction is as follows:  $B$  is played when  $\{M, B\}$  is expected (restrict attention to information sets containing 2 actions for simplicity - hence, no other set is possible to justify  $B$  because of  $\frac{1}{3}$ -stability),  $M$  is played when another set of 2 actions is expected (which must necessarily be  $\{U, B\}$ ) and  $U$  is played when  $\{U, M\}$  is expected. Consider the set  $\Omega$  of "sunspot states" and the information partition of each player. At a correlated equilibrium, for every player, at an element in this partition where  $U$  (respectively  $M$ , respectively  $B$ ) is played, we want that the other player plays the following actions with positive probability (and only these actions):  $U, M$  (respectively  $U, B$ , respectively  $M, B$ ). The simplest set  $\Omega$  consistent with this contains 6 states, information partition are

- For player *Row*:  $\{1, 2\}, \{4, 5\}, \{3, 6\}$  and the respective equilibrium actions must be  $U, M, B$
- For player *Col*:  $\{1, 4\}, \{2, 3\}, \{5, 6\}$  and the respective equilibrium actions must be  $U, M, B$
- In every element of the partition, the beliefs are of the form  $(1/2, 1/2)$  so that  $p = 1/3 < 1/2$ ; hence with a uniform common prior the belief potential is  $\frac{1}{2}$ .
- One can also define the prior probability as follows:  $\pi$  on state 1,  $\frac{1-\pi}{5}$  on every other state. The correlated equilibrium described above is still a correlated equilibrium with this prior (state 1 is the state where  $(U, U)$  is played). If  $\pi = 2/3$ , then the belief potential is  $1/3$ . We therefore have a correlated equilibrium with belief potential  $\frac{1}{3}$  and  $(U, U)$  is played with probability strictly greater than  $\frac{1}{3}$  but other actions are played as well. Ex-ante, it is common knowledge that  $(U, U)$  is played with probability strictly greater than  $\frac{1}{3}$ , but the interim information provided on the other player's action is such that  $(U, U)$  is not necessarily played (at the interim stage, it is not always true that  $(U, U)$  is  $1/3$ -believed).■

Finally, we examine the link between  $p$ -BR, iterated  $p$ -BR (Tercieux, 2006) and  $p$ -stability. We show that although iterated  $p$ -BR sets and  $p$ -stability are equivalent in finite games,  $p$ -stability is adapted to game with continuous set of actions, while iterated  $p$ -stability is not.

We begin with a short summary of the key notation in Tercieux (2006) relevant for our purposes. Consider a game  $G$ : 2 players, set of actions  $A_i$  compact in  $R^L$ ; a game  $G[S]$ : game  $G$  with set of actions restricted to  $S$ ;  $\Pi^p(S_{-i})$   $p$ -belief on  $S_{-i}$ ,  $\Lambda_i(S_{-i}, p)$  best response of player  $i$  to  $p$ -belief on  $S_{-i}$ ,  $\Lambda(S, p)$  product of BR sets;  $p$ -BR set:  $S$  such that  $\Lambda(S, p) \subset S$  (if everyone  $p$ believes in  $S$ , then the outcome is in  $S$ );  $p$ -MBR set:  $p$ -BR set with no proper subset being a  $p$ -BR set; Iterated  $p$ -BR set: there exists a decreasing sequence  $S^0, \dots, S^n$  such that  $S^0 = A$  and  $S^{l+1}$  is a  $p$ -BR set in  $G[S^l]$  and  $S^n = S$ . Tercieux shows that  $p$ -BR sets are iterated  $p$ -BR sets (with  $S^0 = A$  and  $S^1 = S$ ). The converse is not necessarily true. Some remarks: (1) A strategy profile is a 1-MBR set iff it is a strict Nash, minimal curb set (Basu and Weibull (1991)). In a 2x2 game, risk dominance = 1/2-MBR set. An action profile is  $p$ dominant equilibrium iff it is a  $p$ -MBR set. (2) Theorem 1 in Tercieux (2006) shows the existence of  $p$ -BR sets. (3) Although Tercieux doesn't address this point directly, how do players coordinate on a  $p$ -BR set? If it is common knowledge that everyone  $p$ believes in a  $p$ -BR set  $S$  and Common knowledge of rationality, then everyone 1believes in  $S$ . Hence, a  $p$ -BR set is a self-fulfilling set (= a 1-BR set) where the self-fulfilling property is robust to  $p < 1$  (some strategic uncertainty). (4) Recall that  $p$ -dominance is particularly interesting for  $p = 1/2$ . If an equilibrium is  $p$ dominant, then there is no  $(1-p)$ -BR set not containing this equilibrium. Furthermore, if there is a  $p$ -BR set  $S$ , then there is no  $(1-p)$ -BR set disjoint from  $S$

**Result 1.  $p$ -MBR implies  $p$ -stability; the converse is not true.**

**Proof.** Common knowledge of rationality and  $p$ -belief on a Nash equilibrium  $\tau^*$  implies common knowledge that the outcome is in the  $p$ -MBR including  $\tau^*$ . Furthermore, " $\{\tau^*\}$  is a  $p$ -MBR" implies  $p$ -stability (" $\{\tau^*\}$  is a  $p$ -MBR" means  $p$ -dominance, that is:  $p$ -stability in 1 step). However, the converse implication is not true. Example 2.5 in Tercieux (2006) contains a  $p$ -stable equilibrium that is not a  $p$ -MBR set. This is intuitive:  $p$ -BR is only one step of elimination of non best-response, while  $p$ -stab is an iterative process).■

**Result 2.** In a finite game, for a strict Nash  $\tau^*$ ,  $p$ -stability of  $\tau^*$  is equivalent to the statement " $\{\tau^*\}$  is an iterated  $p$ -BR set".

**Remark about the proof.** The proof that  $\{\tau^*\}$  is an iterated  $p$ -BR set is straightforward. To prove the converse implication, we need an auxiliary result showing that, whenever  $\tau^*$  is an iterated  $p$ -BR set, it can be associated with a "canonical" sequence of sets.

**Proof.** (i) Consider a strict Nash  $\tau^*$  that is  $p$ -stable. We have  $S_p^N = \{\tau^*\}$  for some  $N$  (denote  $N$  the smallest integer with this property). We show that  $\{\tau^*\}$  is an iterated  $p$ -BR set. By definition,  $\tau^*$  is a  $p$ -MBR in  $S_p^{N-1}$  (to be rigorous, one needs to distinguish between the  $S_p^n$  of the paper that correspond to mixed strategy sets and the  $S_p^n$  used here that are implicitly pure strategy sets - we don't do this here, this should not be a problem as it seems obvious that the  $S_p^n$  in mixed strategy is the set of all the mixed strategy based on the pure strategy in  $S_p^n$ ).  $S_p^{N-1}$  is a  $p$ -BR in  $S_p^{N-2}$  (as  $\tau^* \in S_p^{N-1}$ ). We now show that it is a  $p$ -MBR. Assume there is a  $p$ -MBR  $S''$  in  $S_p^{N-2}$  including  $\tau^*$  and included in  $S_p^{N-1}$ . Observe that (1) believes  $p$  on  $\tau^*$  and  $(1-p)$  on  $S_p^{N-2}$  implies action in  $S_p^{N-1}$ , (2) believes  $p$  on  $S''$  and  $(1-p)$  on  $S_p^{N-2}$  implies action in  $S''$ , (3) then  $S_p^{N-1} \subset S''$ , then  $S_p^{N-1} = S''$ . We have shown that  $S_p^{N-1}$  is the  $p$ -MBR in  $S_p^{N-2}$  including  $\tau^*$ . We iterate the argument:  $S_p^{N-2}$  is a  $p$ -BR set in  $S^{N-3}$ . We show that it is a  $p$ -MBR. Assume there is a  $p$ -MBR  $S''$  in  $S_p^{N-3}$  including  $\tau^*$  and included in  $S_p^{N-2}$ . Observe that (1) believes  $p$  on  $\tau^*$  and  $(1-p)$  on  $S_p^{N-3}$  implies action in  $S_p^{N-2}$ , (2) believes  $p$  on  $S''$  and  $(1-p)$  on  $S_p^{N-3}$  implies action in  $S''$ , (3) then  $S_p^{N-2} \subset S''$ , then  $S_p^{N-2} = S''$ . We have shown that  $S_p^{N-2}$  is the  $p$ -MBR in  $S_p^{N-3}$  including  $\tau^*$ . Iterating the argument shows that  $\{\tau^*\}$  is an iterated  $p$ -BR set (associated with  $S_p^n$ ). This shows the direct implication.

(ii) Define the sequence  $MBR_n$  as follows:  $MBR_0 = S$ , and for every  $n$ ,  $MBR_{n+1}$  is the  $p$ -MBR including  $\tau^*$  in the game with a restricted strategy set  $MBR_n$ . Notice that the above proof of the direct implication shows that, whenever  $\tau^*$  is  $p$ -stable, the sequence  $S_p^n$  is the sequence  $MBR_n$ . Assume that  $\{\tau^*\}$  is an iterated  $p$ -BR set associated with a sequence  $BR_n$  (denote  $\{\tau^*\} = BR_N$ ). We compare  $BR_n$  to  $MBR_n$ . By definition,  $MBR_1 \subset BR_1$ . Observe that (a) believes  $p$  on  $MBR_2$  and  $(1-p)$  on  $MBR_1$  implies action in  $MBR_2$ , (b) believes  $p$  on  $BR_2$  and  $(1-p)$  on  $BR_1$  implies action in  $BR_2$ , (c) As  $MBR_1 \subset BR_1$ ,  $BR_2$  is a  $p$ -BR set in  $MBR_1$ . Hence,  $MBR_2 \subset BR_2$  (iterating the argument shows that  $MBR_n \subset BR_n$  and  $MBR_N = \{\tau^*\}$ ). We have shown

that  $\{\tau^*\}$  is an iterated  $p$ -BR set associated with a sequence  $MBR_n$ . We now show that  $\tau^*$  is  $p$ -stable. We compare  $S_p^n$  and  $MBR_n$ . So, (a) believes  $p$  on  $\tau^*$  and  $(1-p)$  on  $S$  implies action in  $S_p^1$ , (b) believes  $p$  on  $MBR_1$  and  $(1-p)$  on  $S$  implies action in  $MBR_1$ , (c) then  $S_p^1 \subset MBR_1$ . We iterate the argument. Assume that  $S_p^n \subset MBR_n$ . Then: (a) believes  $p$  on  $\tau^*$  and  $(1-p)$  on  $S_p^n$  implies action in  $S_p^{n+1}$ , (b) believes  $p$  on  $MBR_{n+1}$  and  $(1-p)$  on  $MBR_n$  implies action in  $MBR_{n+1}$ , (c) and  $S_p^{n+1} \subset MBR_{n+1}$ . This shows that  $S_p^N = \{\tau^*\}$ ,  $\tau^*$  is  $p$ -stable. ■

**Result 3. Iterated  $p$ -BR sets are not adapted to games with continuous action sets.**

We demonstrate this via an example. Consider a one-dimensional example with a continuum  $[0, 1]$  of players, the action set is  $[-1, 1]$ , the interior equilibrium is 0 and the BR map is  $a_i = \phi E(\bar{a})$  where  $\bar{a}$  is the average action. We know that the equilibrium is  $p$ -stable for  $\phi(1-p) < 1$ .

**Remark.** There is sometimes corner equilibrium (with equilibrium action being  $-1$  or  $1$ ): this equilibrium isn't considered here.

**Claim 1.** An interval  $[-b, b]$  (for  $b < 1$ ) is a  $p$ -BR set iff  $\phi < 1$  and  $\frac{\phi(1-p)}{1-\phi p} < b$ . Existence of  $p$ -BR sets requires eductive stability ( $\phi < 1$ ).

Notice that the fact that strategic stability is a necessary condition for existence of  $p$ -BR sets is an argument in favor of  $p$ -stability.

**Proof.** The smallest BR to  $p$ -belief in  $[-b, b]$  is  $\phi(-(1-p) - pb)$  and the largest one is  $\phi((1-p) + pb)$ . Then, we have a  $p$ -BR iff  $\phi(1-p) + \phi pb < b$ . That is:  $\phi(1-p) < (1-\phi p)b$ , that is:  $\phi p < 1$  and  $\frac{\phi(1-p)}{1-\phi p} < b$ . This is equivalent to  $\phi < 1$  and  $\frac{\phi(1-p)}{1-\phi p} < b$ . Summing up,  $[-b, b]$  is a  $p$ -BR set when  $\phi < 1$  and  $\frac{\phi(1-p)}{1-\phi p} < b$ . ■

We now characterize when the equilibrium is an iterated  $p$ -BR set.

**Claim 2.** Consider a given  $p$ . The equilibrium is an iterated  $p$ -BR set iff the equilibrium is eductively stable ( $\phi < 1$ ).

Notice here that the fact that the equilibrium is an iterated  $p$ -BR set does not depend on  $p$ , which makes the concept useless. Again, this is an argument in favor of  $p$ -stability.

**Proof.** First consider the case  $\phi > 1$ . Clearly, the equilibrium cannot be an iterated  $p$ -BR set as there is no  $p$ -BR set containing the equilibrium (see Claim 1 above). We now consider the case  $\phi < 1$ . When is an interval  $[-b, b]$  is a  $p$ -BR set in a game restricted to  $[-c, c]$  ( $c > b$ )? The smallest BR to  $p$ -belief in  $[-b, b]$  is  $\phi(-(1-p)c - pb)$  and the largest one is  $\phi((1-p)c + pb)$ . Then, we have a  $p$ -BR iff  $\phi(1-p)c + \phi pb < b$ . That is:  $\frac{\phi c - b}{\phi(c-b)} < p$ . This rewrites  $b > \frac{\phi(1-p)}{1-\phi p}c$ . Define the sequence  $b_n$  by:  $b_1 = \frac{\phi(1-p)}{1-\phi p}$ , and, for every  $n$ ,  $b_{n+1} = \frac{\phi(1-p)}{1-\phi p}b_n$ . We have that  $[-b_{n+1}, b_{n+1}]$  is a  $p$ -BR set in a game restricted to  $[-b_n, b_n]$ . Notice that  $\phi < 1$  implies  $\frac{\phi(1-p)}{1-\phi p} < 1$  and  $b_{n+1} < b_n$ . Hence,  $[-b_{n+1}, b_{n+1}] \subset [-b_n, b_n]$  and the sequence  $[-b_n, b_n]$  converges to  $\{0\}$ . We have just shown that the equilibrium is an iterated  $p$ -BR set for any  $p$ . ■

## 5 Intertemporal trade, expectations coordination and bubbles

In this section, we study the partial consensus outcomes of a standard two period economy to examine the foundations, via belief coordination, of perfect foresight equilibria. We are in a setting where all agents are price-takers and payoffs depend on their own actions and market prices; hence, we adopt a slightly different formalization of a large economy from MasColell (1984). The analysis generalizes and extends Ghosal (2006)'s local stability analysis of a perfect foresight equilibrium: a new solution concept for intertemporal economies is proposed and its links with perfect foresight equilibria is analyzed in a general setting which allows preferences to be non-separable over time.

### 5.1 The Economy

The economy consists of a mass of individuals of measure one, formally, an atomless measure space of individuals,  $\{I, \iota, \mu\}$ , with  $I = [0, 1]$  the set of agents,  $\iota$  the  $\sigma$ -algebra on  $I$  and  $\mu$  an atomless measure defined on  $I$ . Null sets of individuals are systematically ignored throughout the paper. For some arbitrary finite  $K$ -dimensional Euclidian space, an assignment is any function  $\mathbf{g} : I \rightarrow \mathfrak{R}^K$  each coordinate of which is integrable<sup>5</sup>. Trade in this economy is sequential and takes place at two time periods,  $t = 1, 2$ , with  $L_t$ ,  $t = 1, 2$ , commodities traded in the spot commodity markets in period  $t$ <sup>6</sup> and an asset market that opens in the first period. In time period  $t = 1$ , each individual submits commodity demands in the spot commodity markets and asset demands in the asset market. Prices in these markets then adjust to ensure market clearing. In time period  $t = 2$ , each individual submits commodity demands in the spot commodity markets. Prices in these markets then adjust to ensure market clearing.

A commodity bundle is  $x \in \mathfrak{R}_+^{L_1} \times \mathfrak{R}_+^{L_2}$  with  $x_{tl}$  denoting quantities of consumption of commodity  $l$  in period  $t$ . Endowments are  $\mathbf{w} : I \rightarrow \mathfrak{R}_+^{L_1} \times \mathfrak{R}_+^{L_2}$  with  $\mathbf{w} = (\mathbf{w}_1, \mathbf{w}_2)$ , and  $\bar{\mathbf{w}}_t = \int \mathbf{w}_t^i di \gg 0$  for  $t = 1, 2$ . The asset traded in the first period pays off in units of the first commodity traded in the second period and further, it is in zero net supply. Preferences of trader  $i$  are described by a utility function  $u^i : \mathfrak{R}_+^{L_1} \times \mathfrak{R}_+^{L_2} \rightarrow \mathfrak{R}$  such that two assumptions are satisfied: (A1) For each traders  $i \in I$ ,  $u^i$  satisfies strict monotonicity, strict concavity, is twice continuously differentiable on  $\mathfrak{R}_+^{L_1} \times \mathfrak{R}_+^{L_2}$ ; (A2)  $u : I \times \mathfrak{R}_+^{L_1} \times \mathfrak{R}_+^{L_2} \rightarrow \mathfrak{R}$  is measurable and uniformly smooth. The requirement of uniform smoothness in (A2) follows Aumann [2, sections 4 and 10] except that we do not require the utility functions of any trader to be bounded. A consequence of (A2) is that  $u : I \times \mathfrak{R}_+^{L_1} \times \mathfrak{R}_+^{L_2} \rightarrow \mathfrak{R}$  viewed as a map from  $(i, x_1, x_2)$  to real numbers is measurable as a function of  $(i, x_1, x_2)$ . The asset traded in the first period

<sup>5</sup>Throughout this subsection, the bold face type will be used to denote an assignment, with the  $i^{th}$  component of the assignment  $\mathbf{g}$  denoted by  $\mathbf{g}^i$  and the  $k^{th}$  coordinate of the assignment  $\mathbf{g}$  denoted by  $\mathbf{g}_k$ .

<sup>6</sup>Note that when  $L_2 = 1$ , with only one commodity at  $t = 2$ , the coordination problem in second period spot markets studied here disappears. Hence, we assume that  $L_2 \geq 2$ .



pays off in units of the first commodity traded in the second period and further, it is in zero net supply. An allocation is a triple  $(\mathbf{x}_1, \mathbf{y}, \mathbf{x}_2)$  such that  $\mathbf{x}_t^i \in \mathfrak{R}_+^{L_t}$ ,  $t = 1, 2$ , for all  $i \in I$  and  $\mathbf{y}^i \in \mathfrak{R}$ . An allocation is feasible if, in addition,  $\int \mathbf{x}_t^i di = \bar{\mathbf{x}}_t = \bar{\mathbf{w}}_t = \int \mathbf{w}_t^i di$ ,  $t = 1, 2$  and  $\bar{\mathbf{y}} = \int \mathbf{y}^i di = 0$ . An economy is  $\mathbf{E} = \{I, \nu, \mu, (u^i, w^i) : i \in I\}$ .

Prices are  $(\pi_1, q, \pi_2)$  where  $\pi_{tl}$  is the spot commodity market price of commodity  $l$  in period  $t$  and  $q$  is the price of the asset. Normalize prices so that  $\pi_{11} = 1$  and  $\pi_{21} = 1$ . As the utility function of each individual is strongly monotone, without loss of generality it is possible to restrict attention to prices where  $\pi_t \in \mathfrak{R}_{++}^{L_t-1}$ ,  $t = 1, 2$  and  $q \in \mathfrak{R}_{++}$ . Asset payoffs are therefore denoted in the second period numeraire.

At prices  $(\pi_1, q, \pi_2)$  the maximization problem that each individual solves has two stages. In the second stage, at  $t = 2$ , given  $(\pi_1, q, \pi_2, x_1, y)$  each individual solves:

$$\text{Max}_{\{x_2\}} u^i(x_1, x_2) \text{ s.t. } \pi_2 x_2 \leq \pi_2 \mathbf{w}_2^i + y, x_2 \in \mathfrak{R}_+^{L_2}$$

For a solution  $\hat{x}_2^i$  to this maximization problem, let  $v^i(x_1, \pi_2, \pi_2 \mathbf{w}_2^i + y) = u^i(x_1, \hat{x}_2^i(x_1, \pi_2, \pi_2 \mathbf{w}_2^i + y))$ . In the first stage, at  $t = 1$ , given  $(\pi_1, q, \pi_2)$  each individual solves the following maximization problem:

$$\text{Max}_{\{x_1, y\}} v^i(x_1, \pi_2, \pi_2 \mathbf{w}_2^i + y) \text{ s.t. } \pi_1 x_1 + qy \leq \pi_1 \mathbf{w}_1^i, x_1 \in \mathfrak{R}_+^{L_1}$$

Let  $(\hat{x}_1^i, \hat{y}^i)$  denote a solution to this sequential, two-stage maximization problem. Let  $\hat{S}^i(\pi_1, q, \pi_2)$  denote the set of all possible solutions  $(\hat{x}_1^i, \hat{y}^i, \hat{x}_2^i)$  at prices  $(\pi_1, q, \pi_2)$ .

**Definition 1** A Perfect Foresight Equilibrium (PFE) is a vector of prices  $(\hat{\pi}_1, \hat{q}, \hat{\pi}_2)$  and allocations  $(\hat{\mathbf{x}}_1, \hat{\mathbf{y}}, \hat{\mathbf{x}}_2)$  such that (a) at prices  $(\hat{\pi}_1, \hat{q}, \hat{\pi}_2)$ ,  $(\hat{\mathbf{x}}_1^i, \hat{\mathbf{y}}^i, \hat{\mathbf{x}}_2^i) \in \hat{S}^i(\hat{\pi}_1, \hat{q}, \hat{\pi}_2)$ , for all  $i \in I$ , (b)  $\hat{\mathbf{x}}_t \in \mathfrak{R}_{++}^{L_t}$ ,  $t = 1, 2$ , and (c)  $\int \hat{\mathbf{x}}_t^i di = \bar{\mathbf{w}}_t$ ,  $t = 1, 2$  and  $\int \hat{\mathbf{y}}^i di = 0$ .

The corresponding market clearing equations are

$$\int \hat{\mathbf{x}}_1^i(\hat{\pi}_1, \hat{q}, \hat{\pi}_2) di = \bar{\mathbf{w}}_1, \int \hat{\mathbf{y}}^i(\hat{\pi}_1, \hat{q}, \hat{\pi}_2) di = 0, \int \hat{\mathbf{x}}_2^i(\hat{\mathbf{x}}_1^i(\hat{\pi}_1, \hat{q}, \hat{\pi}_2), \hat{\pi}_2, \hat{\mathbf{y}}^i(\hat{\pi}_1, \hat{q}, \hat{\pi}_2)) = \bar{\mathbf{w}}_2$$

Let  $\bar{\mathbf{x}}_t = \int \hat{\mathbf{x}}_t^i di$ ,  $t = 1, 2$  denote mean commodity demand in period  $t$  and  $\bar{\mathbf{y}} = \int \hat{\mathbf{y}}^i di$  denote mean asset demand. Let  $N(\hat{\pi}_1, \hat{q}, \hat{\pi}_2) \subset \mathfrak{R}_{++}^{L_1} \times \mathfrak{R}_{++} \times \mathfrak{R}_{++}^{L_2}$  be a neighborhood of an interior PFE price vector. Then, for all  $(\pi_1, q, \pi_2) \in N(\hat{\pi}_1, \hat{q}, \hat{\pi}_2)$ , the derivatives  $\partial_{\pi_{t'l'}} \bar{\mathbf{x}}_{tl}$ ,  $\partial_{\pi_{t'l'}} \bar{\mathbf{y}}$ ,  $\partial_q \bar{\mathbf{x}}_{tl}$ ,  $\partial_q \bar{\mathbf{y}}$  exist for all  $t, t' = 1, 2$  and all  $l, l' = 1, \dots, L$  and are equal to  $\partial_{\pi_{t'l'}} \bar{\mathbf{x}}_{tl} = \int \partial_{\pi_{t'l'}} \hat{\mathbf{x}}_{tl}^i di$ ,  $\partial_{\pi_{t'l'}} \bar{\mathbf{y}} = \int \partial_{\pi_{t'l'}} \hat{\mathbf{y}}^i di$ ,  $\partial_q \bar{\mathbf{x}}_{tl} = \int \partial_q \hat{\mathbf{x}}_{tl}^i di$ ,  $\partial_q \bar{\mathbf{y}} = \int \partial_q \hat{\mathbf{y}}^i di$ . This follows from the fact that  $\partial_{\pi_{t'l'}} \hat{\mathbf{x}}_{tl}^i$ ,  $\partial_{\pi_{t'l'}} \hat{\mathbf{y}}^i$ ,  $\partial_q \hat{\mathbf{x}}_{tl}^i$ ,  $\partial_q \hat{\mathbf{y}}^i$ , for all  $l, l' = 1, \dots, L_t$ ,  $t, t' = 1, 2$ ,  $i \in I$  are integrally bounded (see, for instance, page 154, Jones (1993)).

After deleting the numeraire commodity in each period, consider the Jacobian of the market clearing equations  $J = \begin{pmatrix} J_{11} & J_{12} \\ J_{21} & J_{22} \end{pmatrix}$  where  $J_{11} = \begin{pmatrix} \partial_{\pi_1} \bar{\mathbf{x}}_1 & \partial_q \bar{\mathbf{x}}_1 \\ \partial_{\pi_1} \bar{\mathbf{y}} & \partial_q \bar{\mathbf{y}} \end{pmatrix}$ ,  $J_{12} = \begin{pmatrix} \partial_{\pi_2} \bar{\mathbf{x}}_1 \\ \partial_{\pi_2} \bar{\mathbf{y}} \end{pmatrix}$ ,  $J_{21} = \begin{pmatrix} \partial_{\pi_1} \bar{\mathbf{x}}_2 & \partial_q \bar{\mathbf{x}}_2 \end{pmatrix}$ ,  $J_{22} = \begin{pmatrix} \partial_{\pi_2} \bar{\mathbf{x}}_2 \end{pmatrix}$  evaluated at the

market clearing prices  $(\hat{\pi}_1, \hat{q}, \hat{\pi}_2) \in \mathfrak{R}_{++}^{L_1-1} \times \mathfrak{R}_{++} \times \mathfrak{R}_{++}^{L_2-1}$ , where the numeraire commodity in each period has been deleted as well. For any assignment  $\mathbf{x}_1, \mathbf{y}$  such that  $\int \mathbf{x}_1^i di = \bar{\mathbf{w}}_1$ ,  $\int \mathbf{y}^i di = 0$ , let  $\partial_{\pi_2} \bar{\mathbf{x}}_2(\mathbf{x}_1, \mathbf{y}) = \int \partial_{\pi_2} \hat{\mathbf{x}}_2^i(\mathbf{x}_1^i, p_2, \mathbf{y}^i)$ .

**Definition 2** (*Regularity, Strong Regularity and Sequential Regularity*)<sup>7</sup> An interior PFE is regular (respectively, strongly regular) if  $J$  is invertible (respectively,  $J_{11}$  and  $J_{22}$  are invertible). It is sequentially regular if, in addition to being regular and strongly regular,  $(\partial_{\pi_2} \bar{\mathbf{x}}_2(\mathbf{x}_1, \mathbf{y}))^{-1}$  exists for all assignment of assets  $\mathbf{y}$  such that  $\mathbf{y}^i = \hat{\mathbf{y}}^i(\pi_1^i, q', \pi_2^i)$  for all  $i \in I$  and some  $(\pi_1^i, q', \pi_2^i) \in N(\hat{\pi}_1, \hat{q}, \hat{\pi}_2, \varepsilon)$ . An economy is regular (respectively, strongly regular and sequentially regular) if all its interior PFE are regular (respectively, strongly regular and sequentially regular).

In a regular economy, each PFE is locally isolated. In a strongly regular economy, in addition, it follows as a consequence of the implicit function theorem that for a given  $(\bar{\pi}_1, \bar{q}) \in N(\hat{\pi}_1, \hat{q}, \varepsilon) \subset \mathfrak{R}_{++}^{L_2-1} \times \mathfrak{R}$ , there is a locally unique second period price  $\pi_2$  that solves  $\int \hat{\mathbf{x}}_2^i(\hat{\mathbf{x}}_1^i(\bar{\pi}_1, \bar{q}, \pi_2), p_2, \hat{\mathbf{y}}^i(\bar{\pi}_1, \bar{q}, p_2)) = \bar{\mathbf{w}}_2$ ; further, it is a continuous function of  $(\bar{\pi}_1, \bar{q}) \in N(\hat{\pi}_1, \hat{q}, \varepsilon)$ . Moreover, in a sequentially regular economy, again as a consequence of the implicit function theorem, for all  $(\pi_1^i, q', \pi_2^i) \in N(\hat{\pi}_1, \hat{q}, \hat{\pi}_2, \varepsilon)$  with  $\int \hat{\mathbf{x}}_1^i(\pi_1^i, q', \pi_2^i) di = \bar{\mathbf{w}}_1$  and  $\int \hat{\mathbf{y}}^i(\pi_1^i, q', \pi_2^i) di = 0$ , there exists a unique second period price  $\pi_2$  that solves  $\int \hat{\mathbf{x}}_2^i(\pi_1^i, q', \pi_2^i), \pi_2, \hat{\mathbf{y}}^i(\pi_1^i, q', \pi_2^i)) = \bar{\mathbf{w}}_2$ ; further, it is a continuous function of  $(\pi_1^i, q', \pi_2^i) \in N(\hat{\pi}_1, \hat{q}, \hat{\pi}_2, \varepsilon)$ .

## 5.2 Partial consensus outcomes, Perfect Foresight Equilibria and Bubbles

Our starting point is the assumption that it is common knowledge that individuals have expectations over future prices whose support is some set  $\Pi_2^0 \subset \mathfrak{R}_{++}^{L_2-1}$ . Let  $\tilde{\pi}_2 \in \Pi_2^0$ . A  $p$ -belief<sup>8</sup> puts a weight  $r$  on  $\tilde{\pi}_2$  and a weight  $1 - p$  on  $\pi_2^{e,i} \in \Pi_2^0$  (not assumed to be common knowledge). The interpretation is that there is partial consensus (where  $p$  measures the degree of consensus) on  $\tilde{\pi}_2$ . Let  $\mathbf{f} : I \rightarrow \Pi_2^0 \subset \mathfrak{R}_{++}^{L_2-1}$ : an assignment of expectations where  $\mathbf{f}^i = \pi_2^{e,i}$ . Let  $\hat{\mathbf{x}}_2^i(x_1, \tilde{\pi}_2, \tilde{\pi}_2 \mathbf{w}_2^i + y)$  denote the (unique) solution to

$$\text{Max}_{\{x_2\}} u^i(x_1, x_2) \text{ s.t. } \tilde{\pi}_2 x_2 \leq \tilde{\pi}_2 \mathbf{w}_2^i + y, x_2 \in \mathfrak{R}_+^{L_2}$$

Let  $\hat{\mathbf{x}}_2^i(x_1, \mathbf{f}^i, \mathbf{f}^i \mathbf{w}_2^i + y)$  denote the (unique) solution to

$$\text{Max}_{\{x_2\}} u^i(x_1, x_2) \text{ s.t. } \mathbf{f}^i x_2 \leq \mathbf{f}^i \mathbf{w}_2^i + y, x_2 \in \mathfrak{R}_+^{L_2}$$

For each  $p$ -belief, there is an associated lottery over period 2 consumption  $l_p^i$ , with probability  $p$  on  $\hat{\mathbf{x}}_2^i(x_1, \tilde{\pi}_2, \tilde{\pi}_2 \mathbf{w}_2^i + y)$  and with probability  $1 - p$  on

<sup>7</sup>Sequential regularity was introduced by Balasko (1994).

<sup>8</sup>The slight change in notation is made for ease of exposition.

$\hat{\mathbf{x}}_2^i(x_1, \mathbf{f}^i, \mathbf{f}^i \mathbf{w}_2^i + y)$ . Let  $\mathbf{l}_p$  denote an assignment of lotteries. At  $t = 1$ , given  $(\pi_1, q)$  and  $l_r^i$  an individual solves:

$$\text{Max}_{\{x_1, y\}} pv^i(x_1, \tilde{\pi}_2, \tilde{\pi}_2 \mathbf{w}_2^i + y) + (1-p)v^i(x_1, \mathbf{f}^i, \mathbf{f}^i \mathbf{w}_2^i + y) \text{ s.t. } \pi_1 x_1 + qy \leq \pi_1 \mathbf{w}_1^i, x_1 \in \mathfrak{R}_+^{L_1}.$$

Let  $\hat{\mathbf{x}}_1^i(\pi_1, q, \mathbf{l}_p^i)$ ,  $\hat{\mathbf{y}}^i(\pi_1', q', \mathbf{l}_p^i)$  denote a solution to the preceding sequential, two-stage maximization problem.

For a fixed  $\mathbf{l}_p$ ,  $(\pi_1', q')$  is a *period 1 equilibrium* if and only if  $\int \hat{\mathbf{x}}_1^i(\pi_1', q', \mathbf{l}_p^i) di = \int \mathbf{w}_1^i di$  and  $\int \hat{\mathbf{y}}^i(\pi_1', q', \mathbf{l}_p^i) di = 0$ . Let  $\hat{E}_1(\mathbf{l}_p)$  denote the set of such equilibria. For a pair  $(\pi_1, q)$  first period prices, with a mild abuse of notation, the second period price  $\pi_2'$  is a *period 2 equilibrium* if and only if  $\int \hat{\mathbf{x}}_2^i(\hat{\mathbf{x}}_1^i, \pi_2', \hat{\mathbf{y}}^i(\pi_1, q, \mathbf{l}_p^i)) = \int \mathbf{w}_2^i di$ . Let  $\hat{E}_2(\hat{\mathbf{x}}_1, \pi_1, q, \mathbf{l}_p)$  denote the set of such equilibria.

Fix  $\Pi_2^0 \subset \mathfrak{R}_{++}^{L_2-1}$ . For  $n = 1, \dots$ , define

$$\Pi_{2,p}^n = \left\{ \begin{array}{l} \pi_2 \in \mathfrak{R}_{++}^{L_2-1} : \exists \mathbf{l}_p \text{ s.t. } \pi_2 \in \hat{E}_2(\hat{\mathbf{x}}_1, \pi_1, q, \mathbf{l}_p), \\ \text{for some } (\hat{\mathbf{x}}_1, \pi_1, q) \in \hat{E}_1(\mathbf{l}_p) \end{array} \right\} \cap \Pi_{2,p}^{n-1}$$

Obviously,  $\Pi_{2,p}^n \subseteq \Pi_{2,p}^{n-1}$ ,  $n = 1, \dots$

**Rationalizable (second period) price expectations:**  $\tilde{\Pi}_{2,p} = \Pi_{2,p}^\infty$ .

**Partial Consensus outcomes:** For  $0 \leq p < 1$ , given  $\Pi_2^0 \subset \mathfrak{R}_{++}^{L_2-1}$ , a consensus outcome is a triple  $(\tilde{\pi}_1, \tilde{q}, \tilde{\pi}_2)$  and an allocation  $(\tilde{\mathbf{x}}_1, \tilde{\mathbf{y}}, \tilde{\mathbf{x}}_2)$  such that  $(\tilde{\mathbf{x}}_1, \tilde{\pi}_1, \tilde{q}) \in \hat{E}_1(\mathbf{l}_p)$  for some  $\mathbf{l}_p$  on  $\tilde{\Pi}_{2,p}$  with probability at least  $p$  on  $\tilde{\pi}_2 \in \tilde{\Pi}_{2,p}$ .

**Proposition 6.** Consider a PFE vector of prices  $(\hat{\pi}_1, \hat{q}, \hat{\pi}_2)$ . Suppose  $u^i(x_1, x_2) = u_1^i(x_2) + u_2(x_2)$  where  $u_2(\cdot)$  homothetic for all  $i \in I$ . Then, for any  $\Pi_2^0 \subset \mathfrak{R}_{++}^{L_2-1}$ , such that  $\hat{\pi}_2 \in \Pi_2^0$ ,  $\tilde{\Pi}_{2,p} = \{\hat{\pi}_2\}$  for all  $0 \leq p < 1$ .

**Proof.** As market clearing in both periods is common knowledge  $\int \hat{\mathbf{y}}^i di = 0$ , and therefore,  $\int d\hat{\mathbf{y}}^i di = 0$ . As  $u_2^i(x_2)$  is identical and homothetic for all  $i \in I$ ,

$$\partial_y \hat{\mathbf{x}}_2^i(\hat{\pi}_2, \hat{\mathbf{y}}^i(\hat{\pi}_1, \hat{q}, \hat{\pi}_2)) = \partial_y \hat{\mathbf{x}}_2^j(\hat{\pi}_2, \hat{\mathbf{y}}^j(\hat{\pi}_1, \hat{q}, \hat{\pi}_2))$$

for all  $i, j \in I$ , and therefore,  $\int \partial_y \hat{\mathbf{x}}_2^i d\hat{\mathbf{y}}^i di = \partial_y \hat{\mathbf{x}}_2^i \int d\hat{\mathbf{y}}^i di = 0$  so that as long as  $\hat{\pi}_2 \in \Pi_2^0$ ,  $\hat{E}_2(\mathbf{x}_1, \pi_1, q, \mathbf{l}_p) = \{\hat{\pi}_2\}$  for all  $(\pi_1, q) \in \hat{E}_1(\mathbf{l}_p)$ ,  $0 \leq p < 1$ . Therefore,  $\tilde{\Pi}_{2,p} = \{\hat{\pi}_2\}$ . ■

A partial consensus outcomes reduces to a PFE when all individuals are able to continue eliminating prices till  $\hat{\pi}_2$  is the only element in  $\tilde{\Pi}_{2,p}$ . When preferences are additively separable in consumption across the two time periods, a change in the asset holdings in period 1 amounts to a redistribution of revenue in period 2 spot markets. When preferences over consumption in period 2 spot markets are identical and homothetic, a redistribution of revenue will have no impact on period 2 spot prices. Therefore, as long the PFE second period price  $\hat{\pi}_2$  is in  $\Pi_2^0$ , it is the only price vector consistent with market clearing in period 2 irrespective of what second price expectations individuals started out with: it is the *unique* rationalizable second price expectation.

**Proposition 7.** Consider a sequentially regular PFE vector of prices  $(\hat{\pi}_1, \hat{q}, \hat{\pi}_2)$ . Whenever the degree of consensus  $p$  on  $\hat{\pi}_2 \in \Pi_2^0$  is higher than a critical threshold value  $\bar{p} < 1$ , the PFE is the unique rationalizable outcome.

**Proof.** When the PFE is sequentially regular, then the admissibility condition required for Proposition 2 is satisfied. By relabelling variables appropriately, the result is an immediate consequence of Proposition 2. ■

Fix a strongly regular  $(\hat{\pi}_1, \hat{q}, \hat{\pi}_2)$ . Clearly, as long as  $\hat{\pi}_2 \in \Pi_2^0$ , there is a partial consensus outcome that is Pareto optimal. However, by Proposition 1, when  $\tilde{\Pi}_{2,p}$  has a non-empty interior in  $\mathfrak{R}_{++}^{L_2-1}$ , it also contains partial consensus outcomes are distinct from a PFE such that the associated allocations aren't Pareto optimal. Evidently, the marginal rates of substitution will be equalized across individuals in spot markets within a time period but not across time periods. Moreover, at a strongly regular  $(\hat{\pi}_1, \hat{q}, \hat{\pi}_2)$ , the associated partial consensus asset price  $\tilde{q} \neq \hat{q}$ : hence, there is an asset price bubble.

In what follows, starting from a fixed strongly regular  $(\hat{\pi}_1, \hat{q}, \hat{\pi}_2)$  and a  $\Pi_2^0$  such that  $\hat{\pi}_2 \in \Pi_2^0$ , a local characterization of the as set of sufficient conditions that ensures the existence of a  $\tilde{\Pi}_{2,p}$  that has a non-empty interior in  $\mathfrak{R}_{++}^{L_2-1}$  is carried out for  $p$  small enough. Suppose  $\Pi_2^0 = N(\hat{\pi}_2, \varepsilon)$ ,  $\Pi_2^0 \neq \{\hat{\pi}_2\}$  with  $N(\hat{\pi}_2, \varepsilon) \subset \mathfrak{R}_{++}^{L_2-1}$  a neighborhood around  $\hat{\pi}_2$ , the second period component of a PFE vector of prices  $(\hat{\pi}_1, \hat{q}, \hat{\pi}_2)$ . Let  $\|\cdot\|$  be a monotone vector norm<sup>9</sup> on  $\mathfrak{R}^{L_2-1}$ . Let  $S(\varepsilon) = \{z \in \mathfrak{R}_{++}^{L_2-1} : \|z - \hat{\pi}_2\| = \varepsilon, \varepsilon > 0\}$ . Let  $B(\bar{\varepsilon}) = \{x \in \mathfrak{R}_{++}^{L_2-1} : \|x - \hat{\pi}_2\| < \bar{\varepsilon}\}$ .

As the economy is sequentially regular it is possible to provide a local characterization of the map from each assignment of expectations  $\mathbf{f} : I \rightarrow \Pi_2^0$  (remember  $p = 0$ ) to a market clearing price in the second period  $\pi_2$  in the vicinity of a PFE.

**Proposition 8:** Fix a sequentially regular PFE  $(\hat{\pi}_1, \hat{q}, \hat{\pi}_2)$ . There exists a neighborhood  $N(\hat{\pi}_2, \varepsilon)$  of  $\hat{\pi}_2$  and matrices  $M^i$  for each  $i \in I$  such that for each assignment of expectations  $\mathbf{f} : I \rightarrow \Pi_2^0$ ,  $\Pi_2^0 = N(\hat{\pi}_2, \varepsilon)$ ,  $d\pi_2 = \int \mathbf{M}^i d\mathbf{f}^i di$  where  $d\pi_2 = (\pi_2 - \hat{\pi}_2)$  and  $d\mathbf{f}^i = (\mathbf{f}^i - \hat{\pi}_2)$ . Moreover  $\tilde{\Pi}_{2,0} = \{\pi_2 \in N(\hat{\pi}_2, \varepsilon) : d\pi_2 = \int \mathbf{M}^i d\mathbf{f}^i di, \text{ for some } \mathbf{f} : I \rightarrow \tilde{\Pi}_{2,0}\}$ . If there exists  $\bar{\varepsilon} > 0$  such that  $\Pi_2^0 \subseteq N(\hat{\pi}_2, \bar{\varepsilon})$ , (a)  $\|\int \mathbf{M}^i d\mathbf{v}^i di\| < \bar{\varepsilon}$ , for all assignment of expectations  $\mathbf{v} : I \rightarrow S(\bar{\varepsilon})$ , then  $\tilde{\Pi}_{2,p} = \{\hat{\pi}_2\}$ , (b) if  $\|\int \mathbf{M}^i d\mathbf{v}^i di\| > \varepsilon$ , for all assignment of expectations  $\mathbf{v} : I \rightarrow S(\varepsilon)$ , for each  $\varepsilon \leq \bar{\varepsilon}$ ,  $N(\hat{\pi}_2, \bar{\varepsilon}) \subseteq \tilde{\Pi}_{2,p}$  for  $p < \bar{p}$ , for some  $\bar{p} > 0$ . Moreover, the condition that  $\|\int \mathbf{M}^i d\mathbf{v}^i di\| < \bar{\varepsilon}$  is invariant to the choice of the second period numeraire.

**Proof.** See appendix. ■

Heuristically, what the preceding proposition shows is that when redistributions of revenue in second period markets change or when redistributions of commodities in period 1 change second period spot market prices *and* there is lack of consensus over second period prices (so that beliefs over second period prices are heterogeneous to a sufficient degree), then an asset price bubble exists.

To interpret the condition under which, locally, partial consensus outcomes

<sup>9</sup> For any  $x \in \mathfrak{R}^K$ , let  $|x| = (|x_1|, \dots, |x_K|)$ . It follows that  $|x| \geq |y|$  if and only if  $|x_l| \geq |y_l|$  for all  $l = 1, \dots, K$ . A vector norm,  $\|\cdot\|$  on  $\mathfrak{R}^K$  is monotone if and only if  $|x| \geq |y| \implies \|x\| \geq \|y\|$ . All  $l_p$  norms, including the euclidean norm, are monotone. However, (see Horn and Johnson (1985)) the following vector norm on  $\mathfrak{R}^K$ ,  $\|x\| = |x_1 - x_2| + \sum_{l' \neq 1} |x_{l'}|$ , is not monotone.

coincide with perfect foresight equilibria, it is useful to consider the special case with homogeneous expectations. In this case, the condition in part (ii) of Proposition 1 reduces to the condition that  $\|\bar{\mathbf{M}}(\pi^e - \hat{\pi}_2)\| < \bar{\varepsilon}$  whenever  $\|\pi^e - \hat{\pi}_2\| = \bar{\varepsilon}$  which is equivalent to requiring that  $\|\bar{\mathbf{M}}\| < 1$ . This is nothing but the requirement that the map that goes from expectations of second period prices to a market clearing second period price is locally a contraction map.

What conditions ensure that redistributions of revenue in second period markets change or when redistributions of commodities in period 1 change second period spot market prices? The following proposition provides a (negative) answer:

**Proposition 9.** Consider a sequentially regular PFE vector of prices  $(\hat{\pi}_1, \hat{q}, \hat{\pi}_2)$ . If there exists  $\tilde{\varepsilon} > 0$  such that either (i)

$$\max \left\{ \left\| \partial_y \hat{\mathbf{x}}_2^i(\hat{\mathbf{x}}_1^i, \hat{\pi}_2, \hat{\mathbf{y}}^i) - \partial_y \hat{\mathbf{x}}_2^j(\hat{\mathbf{x}}_1^i, \hat{\pi}_2, \hat{\mathbf{y}}^i) \right\|, \left\| \partial_{x_1} \hat{\mathbf{x}}_2^i(\hat{\mathbf{x}}_1^i, \hat{\pi}_2, \hat{\mathbf{y}}^i) - \partial_{x_1} \hat{\mathbf{x}}_2^j(\hat{\mathbf{x}}_1^i, \hat{\pi}_2, \hat{\mathbf{y}}^i) \right\| \right\} < \tilde{\varepsilon}$$

for all  $i, j \in I$ , or (ii)  $\max \{ \|\partial_{\pi_2} \hat{\mathbf{y}}^i\|, \|\partial_{\pi_2} \hat{\mathbf{x}}_1^i\| \} < \tilde{\varepsilon}$  for all  $i \in I$ . Then there exists a neighborhood  $N(\hat{\pi}_2, \varepsilon)$  of  $\hat{\pi}_2$  and  $\Pi_2^0 \subset N(\hat{\pi}_2, \varepsilon)$  such that  $\tilde{\Pi}_{2,p} = \{\hat{\pi}_2\}$ ,  $0 \leq p \leq 1$ .

**Proof.** See appendix. ■

The preceding proposition shows that even with lack of consensus over future prices in a small enough neighborhood of a perfect foresight equilibria, an asset price bubble will not exist if (a) the sensitivity of second period spot commodity prices to a redistribution of revenue in the same period and a redistribution in period 1 endowments is small and (b) the sensitivity of period 1 consumption and asset demands to small changes in expectations of second period prices is small.

Finally, we construct an example with a unique PFE, where with heterogeneity in intertemporal preferences implies the existence of partial consensus outcomes distinct from the PFE outcome because second period spot commodity prices to a redistribution of revenue in the same period and asset demands to small changes in expectations of second period prices, thus demonstrating the robust existence of bubbles.

**Example:** There is a single commodity at  $t = 1$  and two commodities at  $t = 2$ . The preferences of individuals are as follows. There are two types of individuals. Type 1 individuals are those with  $u^i(x) = x_{11} + (1 - \beta) \log x_{21} + \beta \log x_{22}$ ,  $\frac{1}{2} < \beta < 1$  and endowments  $\mathbf{w}^i = (\mathbf{w}_{11}^i, \mathbf{w}_{21}^i, \mathbf{w}_{22}^i) \gg 0$ . Type 2 individuals are those with  $u^i(x) = x_{11} + \beta \log x_{21} + (1 - \beta) \log x_{22}$ , and endowments  $\mathbf{w}^i = (\mathbf{w}_{11}^i, \mathbf{w}_{21}^i, \mathbf{w}_{22}^i) \gg 0$ . Let  $p_2$  denote the price of commodity 2 at  $t = 2$  and let  $\bar{\mathbf{w}}_{tl}$  denote the aggregate endowment of commodity  $l$  at time  $t$ . Both type 1 and type 2 individuals have equal measure. Let  $\mathbf{w}'_{21}$  denote the aggregate endowments of the numeraire commodity at  $t = 2$  for type 1 individuals and  $\mathbf{w}''_{21}$  denote the aggregate endowments of the numeraire commodity at  $t = 2$  for type 2 individuals. It is assumed that  $\mathbf{w}'_{21} > \mathbf{w}''_{21}$ . Let  $\Pi_2^0 = \left[ \frac{\bar{\mathbf{w}}_{21}}{\bar{\mathbf{w}}_{22}} - \varepsilon, \frac{\bar{\mathbf{w}}_{21}}{\bar{\mathbf{w}}_{22}} + \varepsilon \right]$ . By computation, it is verified that there is a unique perfect foresight equilibrium

configuration of prices  $(\hat{\pi}_2, \hat{q}) = \left( \frac{\bar{\mathbf{w}}_{21}}{\bar{\mathbf{w}}_{22}}, \frac{1}{2\bar{\mathbf{w}}_{21}} \right)$  and

$$M^i = \begin{cases} -\frac{\mathbf{w}'_{21} + \mathbf{w}''_{21}}{\beta\mathbf{w}'_{21} + (1-\beta)\mathbf{w}''_{21}} \left( \frac{1}{2} - \beta \right) & \text{for type 1 individuals} \\ \frac{\mathbf{w}'_{21} + \mathbf{w}''_{21}}{\beta\mathbf{w}'_{21} + (1-\beta)\mathbf{w}''_{21}} \left( \frac{1}{2} - \beta \right) & \text{for type 2 individuals} \end{cases}$$

Let  $\Pi_2^0 = \left[ \frac{\bar{\mathbf{w}}_{21}}{\bar{\mathbf{w}}_{22}} - \varepsilon, \frac{\bar{\mathbf{w}}_{21}}{\bar{\mathbf{w}}_{22}} + \varepsilon \right]$ . We restrict attention to the case where individuals have point expectations over second period prices. Let  $S(\varepsilon) = \{z \in \mathfrak{R}_{++} : |z - \frac{\bar{\mathbf{w}}_{21}}{\bar{\mathbf{w}}_{22}}| = \varepsilon\}$ . Consider the assignment of expectations  $\mathbf{v}_+ : I \rightarrow S(\varepsilon)$  (respectively,  $\mathbf{v}_- : I \rightarrow S(\varepsilon)$ ) where  $\mathbf{v}_+^i = \frac{\bar{\mathbf{w}}_{21}}{\bar{\mathbf{w}}_{22}} - \varepsilon$  (respectively,  $\mathbf{v}_-^i = \frac{\bar{\mathbf{w}}_{21}}{\bar{\mathbf{w}}_{22}} + \varepsilon$ ) for type 1 individuals and  $\mathbf{v}_+^i = \frac{\bar{\mathbf{w}}_{21}}{\bar{\mathbf{w}}_{22}} + \varepsilon$  (respectively,  $\mathbf{v}_-^i = \frac{\bar{\mathbf{w}}_{21}}{\bar{\mathbf{w}}_{22}} - \varepsilon$ ) for type 2 individuals. Then, by computation it is checked that

$$\begin{aligned} \left\| \int \mathbf{M}^i d\mathbf{v}_+^i di \right\| &= \left\| \int \mathbf{M}^i d\mathbf{v}_-^i di \right\| \\ &= \left| \frac{\mathbf{w}'_{21} + \mathbf{w}''_{21}}{\beta\mathbf{w}'_{21} + (1-\beta)\mathbf{w}''_{21}} \left( \frac{1}{2} - \beta \right) \varepsilon \right| > \varepsilon \end{aligned}$$

as long as  $\frac{1}{2} < \beta < 1$  and  $\mathbf{w}'_{21} < \mathbf{w}''_{21}$ . It follows that  $\Pi_{2,0}^n = P_2^0$ ,  $n \geq 1$  and by continuity in  $p$ , there exists  $\tilde{p} > 0$  such that for all  $p < \tilde{p}$ ,  $\Pi_{2,p}^n = P_2^0$ ,  $n \geq 1$ . Hence,  $\tilde{\Pi}_{2,p} = \Pi_2^0$   $p < \tilde{p}$ , for some  $\tilde{p} > 0$ . ■

## 6 Conclusion

In this paper, we have developed a new solution concept that allows for partial consensus about the outcomes of strategic and market interaction and an associated, continuous measure of the degree of stability, via belief coordination, for equilibrium outcomes. In a number of examples we have illustrated and differentiated the properties of our concepts from related notions developed elsewhere. We have examined the foundations of intertemporal trade via belief coordination in a two period economy and show that, under certain conditions, lack of consensus over future prices is consistent with an asset price bubble.

Our contributions are a preliminary step towards understanding how the possibility that non-equilibrium outcomes are approximately self-fulfilling complicate the analysis of strategic interaction and market behavior. In future research, we intend to examine this point in greater detail to obtain new economic insights and their implications for policy.

## References

- [1] Angeletos G.-M., F. Collard and H. Dellas (2018), "Quantifying Confidence", *Econometrica*, Vol. 86, No. 5 (September, 2018), 1689–1726.

- [2] Balasko, Y., (1994): "The expectational stability of Walrasian equilibria", *Journal of Mathematical Economics*, 23, 179-203.
- [3] Cass, D. and K. Shell (1983), "Do Sunspots Matter?", *Journal of Political Economy*, Vol. 91, No. 2 (Apr., 1983), pp. 193-227.
- [4] Chatterji, S. and S. Ghosal (2004), "Local coordination and market equilibria", (with S.Chatterji), *Journal of Economic Theory*, 114, 255-279.
- [5] Dudley, R.M. (1989), *Real Analysis and Probability*, Wadsworth & Brooks/Cole Advanced Books and Software, Wadsworth Inc. Belmont, California.
- [6] Evans, G. and R. Guesnerie (1993), "Rationalizability, Strong Rationality and Expectational Stability", *Games and Economic Behavior*, Special Issue on Learning Dynamics, Vol. 5, October 1993, pp. 632-646.
- [7] Evans, G. W., R. Guesnerie and B. McGough (2019), "Eductive Stability in Real Business Cycle Models", *The Economic Journal*, Doi: 10.1111/eoj.12620
- [8] Farmer, R and K., Platonov (2019), "Animal Spirits in a Monetary Economy", *European Economic Review*, 115, pages 60 – 77.
- [9] Fudenberg, D. and D. K. Levine (1993), "Self-Confirming Equilibrium", *Econometrica*, Vol. 61, No. 3 (May, 1993), pp. 523-545.
- [10] Ghosal, S. (2006), "Intertemporal coordination in two period markets", *Journal of Mathematical Economics*, 43(1), 11-36.
- [11] Guesnerie R. (1992), "An Exploration of the Eductive Justifications of the Rational Expectations Hypothesis", *American Economic Review*, 82,5, 1254-1278.
- [12] Guesnerie, R. (2005), *Assessing Rational Expectations 2: Eductive Stability in Economics*, Cambridge, MA: MIT Press.
- [13] Jones, F., (1993): *Lebesgue integration on Euclidean Space*, Jones and Bartlett Publishers, Boston-London.
- [14] Mas-Colell A. (1984), "On a theorem of Schmeidler", *Journal of Mathematical Economics*, 13, 201-206
- [15] Mertens J.F. and S. Zamir (1985), "Formulation of Bayesian analysis for games with incomplete information", *International Journal of Game Theory*, Volume 14, Number 1 / March, 1985.
- [16] Morris S., R. Rob and H.S. Shin (1995), "p-Dominance and Belief Potential", *Econometrica*, 63,1, 145-157.

- [17] Tercieux O. (2006), "p-Best response set", *Journal of Economic Theory*, 131(1), 45-70.
- [18] Woodford, M. (2013), "Macroeconomic Analysis without the Rational Expectations Hypothesis", *Annual Review of Economics*.

## Appendix Proof of Lemma 1.

(i) For every  $M$ , for every  $\varepsilon > d_{\Delta(A)}(\tau'_a, \hat{\tau}_a)$ , we have

$$\hat{\tau}_a(M) \leq \tau'_a(M^\varepsilon) + \varepsilon,$$

then

$$\begin{aligned} (1-p)\hat{\tau}_a(M) &\leq (1-p)\tau'_a(M^\varepsilon) + (1-p)\varepsilon, \\ \hat{\tau}_a(M) &\leq p\hat{\tau}_a(M) + (1-p)\tau'_a(M^\varepsilon) + (1-p)\varepsilon, \\ \hat{\tau}_a(M) &\leq p\hat{\tau}_a(M^\varepsilon) + (1-p)\tau'_a(M^\varepsilon) + (1-p)\varepsilon. \end{aligned}$$

This implies:  $d_{\Delta(A)}(\tau_a, \hat{\tau}_a) \leq (1-p)\varepsilon$  and

$$d_{\Delta(A)}(\tau_a, \hat{\tau}_a) \leq (1-p)d_{\Delta(A)}(\tau'_a, \hat{\tau}_a)$$

(ii) For a Borel set  $M$  s.t.  $x \in M$ , for every  $\varepsilon$ , we have  $\delta_x(M^\varepsilon) = 1$  and

$$\tau_a(M) \leq \delta_x(M^\varepsilon) + \varepsilon.$$

For a Borel set  $M$  that does not intersect  $S$ , for every  $\varepsilon$ , we have  $\tau_a(M) = 0$  and

$$\tau_a(M) \leq \delta_x(M^\varepsilon) + \varepsilon.$$

Consider now a Borel set  $M$  that does not contain  $x$  and that intersects  $S$ . For every  $\varepsilon > d$ , we have that  $x \in M^\varepsilon$  (consider a  $y$  in  $S \cap M$ ) and  $\delta_x(M^\varepsilon) = 1$  and

$$\tau_a(M) \leq \delta_x(M^\varepsilon) + \varepsilon.$$

(iii) Consider  $\varepsilon > \sup_{\lambda} d_{\Delta(A)}(\tau_\lambda, \nu)$ . For  $f$ -almost every  $\lambda$ , we have: for every  $M$

$$\tau_\lambda(M) \leq \nu(M^\varepsilon) + \varepsilon.$$

Summing over  $\lambda$  gives:

$$\int \tau_\lambda(M) f(d\lambda) \leq \int (\nu(M^\varepsilon) + \varepsilon) f(d\lambda) = \nu(M^\varepsilon) + \varepsilon,$$

as  $(\nu(M^\varepsilon) + \varepsilon)$  does not depend on  $\lambda$  and  $\int f(d\lambda) = 1$ . ■

## Proof of Proposition 2.

There is a neighborhood  $N \subset \Delta(A)$  of  $\tau_a^*$  such that for  $\mu$ -almost every  $u$  in  $U_A$ ,

$$\forall m \in N, d_A(B(u, m), B(u, \tau_a^*)) \leq K d_{\Delta(A)}(m, \tau_a^*). \quad (2)$$



Now for each  $\tau \in T_p$  (with  $\tau = p\tau^* + (1-p)\tau'$ ), we must have that

$$\forall M \subset A, \tau_a(M) = p\tau_a^*(M) + (1-p)\tau_a'(M),$$

It straightforwardly follows from Lemma 1(i) that

$$d_{\Delta(A)}(\tau_a, \tau_a^*) \leq (1-p)d_{\Delta(A)}(\tau_a', \tau_a^*). \quad (3)$$

As the Prohorov metric is always bounded by 1, we have  $d_{\Delta(A)}(\tau_a', \tau_a^*) \leq 1$  and  $d_{\Delta(A)}(\tau_a, \tau_a^*) \leq 1-p$ . Then, for  $p$  large enough, the following property holds: the inequality (2) holds for the marginal  $\tau_a$  of any distribution  $\tau$  in  $T_p$ . From now on, we consider  $p$  such that this property holds. Define the set  $A^n(u) \subset A$  of actions that are best responses of  $u$  to a distribution of actions  $\tau_a$  that is the marginal on  $A$  of some  $\tau \in S_p^{n-1}$ .  $\phi(S_p^{n-1})$  contains the distributions  $\tau \in T$  such that, for  $\mu$ -almost every  $u$ ,  $\tau(A^n(u)|u) = 1$ . For  $\mu$ -almost every  $u$ , for every  $a$  in  $A^n(u)$ ,  $a$  writes  $B(u, \tau_a)$  for some  $\tau$  in  $S_p^{n-1}$ . Inequality (2) writes:

$$d_A(a, B(u, \tau_a^*)) \leq K d_{\Delta(A)}(\tau_a, \tau_a^*).$$

As  $\tau \in S_p^{n-1} = \phi(S_p^{n-2}) \cap T_p$ ,  $\tau = p\tau^* + (1-p)\tau'$  for some  $\tau' \in \phi(S_p^{n-2})$ . Inequality (3) implies

$$d_A(a, B(u, \tau_a^*)) \leq K(1-p)d_{\Delta(A)}(\tau_a', \tau_a^*).$$

Denote  $R_A^n(u) = \sup_{a \in A^n(u)} d_A(a, B(u, \tau_a^*))$  (this is the radius of the smallest ball containing  $A^n(u)$  and centered on  $\tau_a^*$ ). We have

$$R_A^n(u) \leq K(1-p)d_{\Delta(A)}(\tau_a', \tau_a^*).$$

Denote  $R_A^n = \sup \text{ess}_{u \in U_A} R_A^n(u)$  for every  $n$ . We have

$$R_A^n \leq K(1-p)d_{\Delta(A)}(\tau_a', \tau_a^*). \quad (4)$$

Consider now that by definition, for every  $M$ ,  $\tau_a^*(M) = \int \tau^*(M|u) \mu(du)$ . By admissibility of  $\tau^*$ , for  $\mu$ -almost every  $u$ , the conditional distribution  $\tau^*(\cdot|u)$  is a Dirac measure  $\delta_{B(\tau_a^*, u)}$  on the equilibrium action of  $u$  (denoted  $B(\tau_a^*, u)$ ). By Lemma 1(ii), for every Dirac measure centered on  $x \in A$ ,

$$d_{\Delta(A)}(\tau_a', \delta_x) \leq \sup_{y \in S} d_A(x, y), \quad (5)$$

where  $S$  is the support of  $\tau_a'$  (the smallest closed set such that  $\tau_a'(S) = 1$ ). Then, we have:

$$d_{\Delta(A)}(\tau_a', \delta_{B(\tau_a^*, u)}) \leq R_A^{n-2}(u)$$

As  $\tau_a^* = \int \delta_{B(\tau_a^*, u)} \mu(du)$ , by Lemma 1(iii),

$$d_{\Delta(A)}(\tau_a', \tau_a^*) \leq \sup_{u \in U_A} \text{ess} d_{\Delta(A)}(\tau_a', \delta_{B(\tau_a^*, u)}) \leq \sup_{u \in U_A} \text{ess} R_A^{n-2}(u) = R_A^{n-2}.$$

From Inequality (4), we have

$$R_A^n \leq K(1-p)R_A^{n-2}.$$

Hence, for  $p$  large enough,  $K(1-p) < 1$  and the sequence of  $R_A^n$  tends to 0, which implies that  $S_p^n$  tends to  $\{\tau^*\}$ . We have shown  $p$ -stability for  $p < 1$  large enough. ■

#### Proof of Proposition 4.

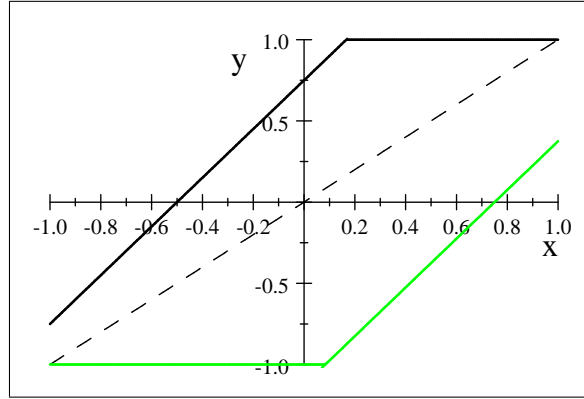
The  $p$ -BR map is  $BR_{\bar{a},p}(a) = \beta(p\bar{a} + (1-p)a)$  if  $\beta(p\bar{a} + (1-p)a) \in [-1, 1]$ , and it is  $BR_{\bar{a},p}(a) = -1$  or  $1$  if  $\beta(p\bar{a} + (1-p)a)$  is smaller than  $-1$  or larger than  $1$ . The sequence  $BR_{\bar{a},p}^n([-1, 1])$  (where  $BR_{\bar{a},p}^n$  denotes the  $n$ -th iterate of  $BR_{\bar{a},p}$ ) converges to a limit set  $BR_{\bar{a},p}^\infty([-1, 1])$  that is the set  $S_{\bar{a},p}^\infty$ . We compute the limit set  $S_{\bar{a},p}^\infty$  in any possible case and check when  $\bar{a} \in S_{\bar{a},p}^\infty$ .

The  $p$ -BR map is converging iff  $|\beta|(1-p) < 1$  (and the limit is  $\frac{\beta p \bar{a}}{1-\beta(1-p)}$ ).

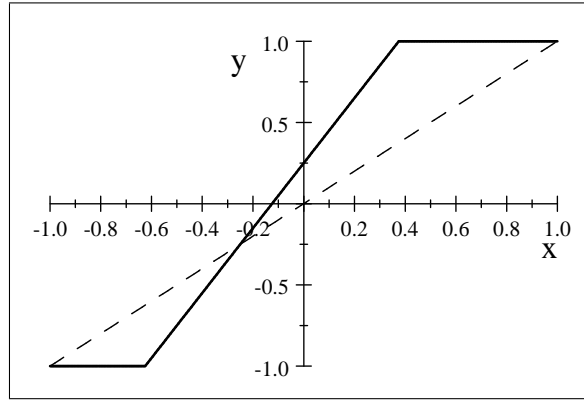
In this case,  $S_{\bar{a},p}^\infty = \left\{ \frac{\beta p \bar{a}}{1-\beta(1-p)} \right\}$  and  $\bar{a} \notin S_{\bar{a},p}^\infty$ :  $\bar{a}$  is not a  $p$ -consensus outcome for a value  $p > 1 - \frac{1}{|\beta|}$ . In the sequel, we assume  $p \leq 1 - \frac{1}{|\beta|}$  (i.e.,  $1 \leq |\beta|(1-p)$ ). Under this assumption, the  $p$ -BR map is not converging to a single point. We study the behavior of the sequence  $BR_{\bar{a},p}^n([-1, 1])$  in the 2 cases  $\beta \geq \frac{1}{1-p}$  and  $\beta \leq -\frac{1}{1-p}$ .

If  $\beta \geq \frac{1}{1-p}$ , then we distinguish between 3 cases:

- If  $\bar{a} > \frac{\beta(1-p)-1}{\beta p}$  ( $BR_{\bar{a},p}(-1) > -1$ ), then  $1$  is the unique fixed point,  $BR_{\bar{a},p}(a) \geq a$  for every  $a$ . Any sequence  $BR_{\bar{a},p}^n(a)$  converges to  $1$ . Hence,  $S_{\bar{a},p}^\infty = \{1\}$ .
- If  $\bar{a} < \frac{1-\beta(1-p)}{\beta p}$  ( $BR_{\bar{a},p}(1) < 1$ ), then  $-1$  is the unique fixed point,  $BR_{\bar{a},p}(a) \leq a$  for every  $a$ . Any sequence  $BR_{\bar{a},p}^n(a)$  converges to  $-1$ . Hence,  $S_{\bar{a},p}^\infty = \{-1\}$ .
- If  $\frac{\beta(1-p)-1}{\beta p} \geq \bar{a} \geq \frac{1-\beta(1-p)}{\beta p}$  ( $BR_{\bar{a},p}(-1) = -1$  and  $BR_{\bar{a},p}(1) = 1$ ), then  $BR_{\bar{a},p}$  admits 3 fixed points (equal to  $-1$ ,  $\frac{\beta p \bar{a}}{1-\beta(1-p)}$  and  $1$ ). For  $a \leq \frac{\beta p \bar{a}}{1-\beta(1-p)}$ , any sequence  $BR_{\bar{a},p}^n(a)$  converges to  $-1$ . For  $a \geq \frac{\beta p \bar{a}}{1-\beta(1-p)}$ , any sequence  $BR_{\bar{a},p}^n(a)$  converges to  $1$ . Hence,  $S_{\bar{a},p}^\infty = [-1, 1]$ .



The cases  $BR_{\bar{a},p}(-1) > -1$  (black) and  $BR_{\bar{a},p}(1) < 1$  (green)



The case  $BR_{\bar{a},p}(-1) = -1$  and  $BR_{\bar{a},p}(1) = 1$

Therefore,  $\bar{a} \in S_{\bar{a},p}^\infty$  iff  $\frac{\beta(1-p)-1}{\beta p} \geq \bar{a} \geq \frac{1-\beta(1-p)}{\beta p}$ . These 2 inequalities rewrite:

$$p \leq \frac{1}{1+\bar{a}} \left(1 - \frac{1}{\beta}\right) \quad \text{and} \quad p \leq \frac{1}{1-\bar{a}} \left(1 - \frac{1}{\beta}\right).$$

This shows the result in the case  $\beta \geq \frac{1}{1-p}$ .

In the case  $\beta \leq -\frac{1}{1-p}$ , we have  $p \in \left[0, 1 + \frac{1}{\beta}\right]$  and  $BR_{\bar{a},p}$  is piecewise linear and decreasing (its slope is either 0 or  $\beta(1-p) < -1$ ). We first consider the case  $\bar{a} \geq 0$ . A preliminary remark is:

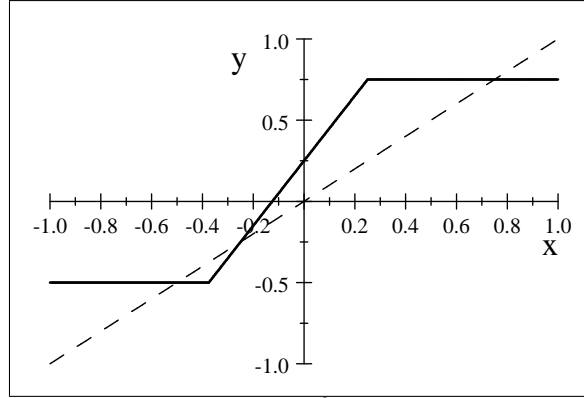
- $BR_{\bar{a},p}(-1) = 1$  iff  $p \leq \frac{1}{1+\bar{a}} \left(1 + \frac{1}{\beta}\right)$
- $BR_{\bar{a},p}(-1) = -1$  iff  $p \geq \frac{1}{1+\bar{a}} \left(1 - \frac{1}{\beta}\right)$
- $BR_{\bar{a},p}(1) = -1$  because  $p \leq \left(1 + \frac{1}{\beta}\right) \leq \frac{1}{1-\bar{a}} \left(1 + \frac{1}{\beta}\right)$  holds true in this case

We have:

1. If  $p \leq \frac{1}{1+\bar{a}} \left(1 + \frac{1}{\beta}\right)$ , then  $BR_{\bar{a},p}(-1) = 1$  and  $BR_{\bar{a},p}(1) = -1$ . Straightforwardly,  $BR_{\bar{a},p}([-1, 1]) = [-1, 1]$  and  $S_{\bar{a},p}^\infty = [-1, 1]$ .
2. If  $\frac{1}{1+\bar{a}} \left(1 + \frac{1}{\beta}\right) < p < \frac{1}{1+\bar{a}} \left(1 - \frac{1}{\beta}\right)$ , then  $-1 < BR_{\bar{a},p}(-1) < 1$  and  $BR_{\bar{a},p}(1) = -1$ . Furthermore,  $BR_{\bar{a},p}(a) = -1$  iff  $a \geq -\frac{1-\beta p \bar{a}}{\beta(1-p)}$ .
3. If  $\frac{1}{1+\bar{a}} \left(1 - \frac{1}{\beta}\right) \leq p \leq 1 + \frac{1}{\beta}$ , then  $BR_{\bar{a},p}(-1) = -1$  and the map  $BR_{\bar{a},p}$  is constant, equal to  $-1$ . Straightforwardly,  $S_{\bar{a},p}^\infty = \{-1\}$ .

In Case 2,  $S_{\bar{a},p}^\infty$  is computed using routine arguments relying on  $BR_{\bar{a},p}^2$ .  $BR_{\bar{a},p}^2$  is increasing and piecewise linear. Since  $\beta^2(1-p)^2 > 1$ ,  $BR_{\bar{a},p}^2$  is constant in the neighborhood of  $-1$  (equal to  $BR_{\bar{a},p}^2(-1)$ ), increasing with slope  $\beta^2(1-p)^2$  for intermediate values of  $a$ , and then constant in the neighborhood of  $1$  (equal to  $BR_{\bar{a},p}^2(1)$ ).

Since  $BR_{\bar{a},p}$  is decreasing, it has a unique fixed point  $a_{fp} = \frac{\beta p \bar{a}}{1-\beta(1-p)}$ . This fixed point  $a_{fp}$  is in  $(-1, 1)$ :  $a_{fp}$  is not on a constant part of  $BR_{\bar{a},p}$ , the slope of  $BR_{\bar{a},p}$  at  $a_{fp}$  is then  $\beta(1-p)$  and the slope of  $BR_{\bar{a},p}^2$  at  $a_{fp}$  is  $\beta^2(1-p)^2$ . Hence,  $BR_{\bar{a},p}^2$  has 3 fixed points:  $BR_{\bar{a},p}^2(-1)$ ,  $a_{fp}$  and  $BR_{\bar{a},p}^2(1)$ , with  $BR_{\bar{a},p}^2(-1)$  and  $a_{fp}$  being negative.



A typical example of  $BR_{\bar{a},p}^2$  with 3 fixed points

A standard argument implies then that the sequence of the iterates  $BR_{\bar{a},p}^n([-1, 1])$  tends to  $[BR_{\bar{a},p}^2(-1), BR_{\bar{a},p}^2(1)]$ . Hence, the set  $S_{\bar{a},p}^\infty$  is  $[BR_{\bar{a},p}^2(-1), BR_{\bar{a},p}^2(1)]$  (with  $BR_{\bar{a},p}^2(1) = BR_{\bar{a},p}(-1)$ ). Since  $\bar{a} \geq 0$ ,  $\bar{a} \in S_{\bar{a},p}^\infty$  iff  $\bar{a} \leq BR_{\bar{a},p}(-1)$ . This condition writes:

$$p \leq \frac{1}{1+\bar{a}} \left(1 + \frac{\bar{a}}{\beta}\right).$$

We now consider the case  $\bar{a} \leq 0$ . A preliminary remark is:

- $BR_{\bar{a},p}(-1) = 1$  because  $p \leq \left(1 + \frac{1}{\beta}\right) \leq \frac{1}{1+\bar{a}} \left(1 + \frac{1}{\beta}\right)$  holds true in this case
- $BR_{\bar{a},p}(1) = -1$  iff  $p \leq \frac{1}{1-\bar{a}} \left(1 + \frac{1}{\beta}\right)$
- $BR_{\bar{a},p}(1) = 1$  iff  $p \geq \frac{1}{1-\bar{a}} \left(1 - \frac{1}{\beta}\right)$

We have:

1. If  $p \leq \frac{1}{1-\bar{a}} \left(1 + \frac{1}{\beta}\right)$ , then  $BR_{\bar{a},p}(-1) = 1$  and  $BR_{\bar{a},p}(1) = -1$ . Straightforwardly,  $BR_{\bar{a},p}([-1, 1]) = [-1, 1]$  and  $S_{\bar{a},p}^\infty = [-1, 1]$ .
2. If  $\frac{1}{1-\bar{a}} \left(1 + \frac{1}{\beta}\right) < p < \frac{1}{1-\bar{a}} \left(1 - \frac{1}{\beta}\right)$ , then  $BR_{\bar{a},p}(-1) = 1$  and  $-1 < BR_{\bar{a},p}(1) < 1$ . Furthermore,  $BR_{\bar{a},p}(a) = 1$  iff  $a \leq \frac{1-\beta p \bar{a}}{\beta(1-p)}$ .
3. If  $\frac{1}{1-\bar{a}} \left(1 - \frac{1}{\beta}\right) \leq p \leq 1 + \frac{1}{\beta}$ , then  $BR_{\bar{a},p}(1) = 1$  and the map  $BR_{\bar{a},p}$  is constant, equal to 1. Straightforwardly,  $S_{\bar{a},p}^\infty = \{1\}$ .

In case 2,  $S_{\bar{a},p}^\infty$  is computed using the same arguments as above.  $BR_{\bar{a},p}^2$  is increasing and piecewise linear, it has 3 fixed points:  $BR_{\bar{a},p}^2(-1)$ ,  $a_{fp}$  and  $BR_{\bar{a},p}^2(1)$ , with  $a_{fp}$  and  $BR_{\bar{a},p}^2(1)$  being positive.  $S_{\bar{a},p}^\infty$  is  $[BR_{\bar{a},p}^2(-1), BR_{\bar{a},p}^2(1)]$  (with  $BR_{\bar{a},p}^2(-1) = BR_{\bar{a},p}(1)$ ). Since  $\bar{a} \leq 0$ ,  $\bar{a} \in S_{\bar{a},p}^\infty$  iff  $\bar{a} \geq BR_{\bar{a},p}(1)$ . This condition writes:

$$p \leq \frac{1}{1-\bar{a}} \left(1 - \frac{\bar{a}}{\beta}\right).$$

Lastly, we summarize the results for any  $\bar{a} \in \mathbb{R}$ . Note that Cases 3 (both when  $\bar{a} \geq 0$  and  $\bar{a} \leq 0$ ) exist iff  $1 + \frac{2}{|\bar{a}|} \leq -\beta$ . If  $1 + \frac{2}{|\bar{a}|} > -\beta$ , then the upper bound for Cases 2 is  $1 + \frac{1}{\beta}$ . Hence,  $\bar{a} \in S_{\bar{a},p}^\infty = [-1, 1]$  when  $p \leq \frac{1}{1+|\bar{a}|} \left(1 + \frac{1}{\beta}\right)$ ;  $\bar{a} \in S_{\bar{a},p}^\infty = [BR_{\bar{a},p}^2(-1), BR_{\bar{a},p}^2(1)]$  when

$$\frac{1}{1+|\bar{a}|} \left(1 + \frac{1}{\beta}\right) < p \leq \min \left( \frac{1}{1+|\bar{a}|} \left(1 - \frac{1}{\beta}\right), \frac{1}{1+|\bar{a}|} \left(1 + \frac{|\bar{a}|}{\beta}\right), 1 + \frac{1}{\beta} \right).$$

For larger values of  $p \leq 1 + \frac{1}{\beta}$ ,  $S_{\bar{a},p}^\infty$  is restricted to one value or  $\bar{a} \notin S_{\bar{a},p}^\infty$ . Hence,  $\bar{a}$  is a  $p$ -consensus outcome with

$$p = \min \left( \frac{1}{1+|\bar{a}|} \left(1 + \frac{|\bar{a}|}{\beta}\right), 1 + \frac{1}{\beta} \right).$$

The minimum is  $1 + \frac{1}{\beta}$  iff  $|\bar{a}| \leq -\frac{1}{\beta}$ . This shows the result in the case  $\beta \leq -\frac{1}{1-p}$ .

### ■ Computations underpinning correlate equilibria in Section 3.1

Given that  $p = 1 + \frac{1}{\beta}$ ,  $\beta < -1$  and  $-1/\beta \geq \bar{a} > 0$ , computations show  $\hat{a}_U < 1$  for  $\varepsilon$  small enough (note that  $\beta\bar{a} + \bar{a} - \bar{a}' < -\bar{a}' < 1$ ) and  $\hat{a}_U > \bar{a}$

follows from  $\bar{a}' < \frac{(1-\varepsilon)\beta}{1-\beta\varepsilon}\bar{a}$ . Hence,  $\hat{a}_U$  is the optimal action for an "uninformed" agent. Computations show  $\hat{a}_I > -1$  (since  $q\bar{a} + (1-q)\bar{a}' < \bar{a} \leq -\frac{1}{\beta}$ ) and  $\hat{a}_I < \bar{a}'$  follows from the lower bound on  $q$ . Hence,  $\hat{a}_I$  is the optimal action for an "informed" agent. Lastly,  $\alpha$  and  $\alpha'$  satisfy the equilibrium conditions. The conditions  $\hat{a}_U \in (\bar{a}, 1)$  and  $\hat{a}_I \in (-1, \bar{a}')$  imply  $\alpha, \alpha' \in (0, 1)$ . ■

A heterogenous prior distribution can be defined.

In the second subcase  $\bar{a} > -1/\beta$ ,  $p = \frac{1}{1+\bar{a}} \left(1 + \frac{\bar{a}}{\beta}\right)$ , for any small enough  $\varepsilon > 0$ , we define the correlated equilibrium as follows. The 2 outcomes are  $\bar{a}$  and some other outcome  $\bar{a}'$ . There are 2 types of agents: Either an agent holds the "uninformed" belief ( $\bar{a}$  with probability  $p - \varepsilon$  and  $\bar{a}'$  with probability  $1 - p + \varepsilon$ ) and plays an action  $\hat{a}_U$  (defined below), or an agent holds the "almost informed" belief ( $\bar{a}$  with probability  $q > p$  and  $\bar{a}'$  with probability  $1 - q$ ) and plays the action  $-1$ . When the outcome is  $\bar{a}$ , a proportion  $\alpha$  of agents (defined below) holds the "uninformed" belief and the remaining proportion holds the "almost informed" belief. When the outcome is  $\bar{a}'$ , the proportion of agents holding the "uninformed" belief is  $\alpha'$  (defined below).

We now write the conditions ensuring that this distribution of actions and beliefs is a correlated equilibrium. The optimal action  $\hat{a}_U$  of an "uninformed" satisfies

$$\hat{a}_U = \beta((p - \varepsilon)\bar{a} + (1 - p + \varepsilon)\bar{a}') \in (-1, 1),$$

and the optimal action of an "almost informed" is  $-1$  if

$$-1 \geq \beta(q\bar{a} + (1 - q)\bar{a}').$$

When the outcome is  $\bar{a}$ , given the distribution of beliefs in the population, the equilibrium condition is

$$\bar{a} = \alpha(-1) + (1 - \alpha)\hat{a}_U,$$

and, when the outcome is  $\bar{a}'$ , the equilibrium condition is

$$\bar{a}' = \alpha'(-1) + (1 - \alpha')\hat{a}_U.$$

An equilibrium is obtained for  $(\alpha, \alpha', q, \bar{a}', \hat{a}_U)$  satisfying

$$\begin{aligned} \bar{a}' &\in \left(-1, \frac{1 - \beta(p - \varepsilon)}{\beta(1 - p + \varepsilon)}\bar{a}\right) \subseteq (-1, 0), \\ q &\in \left(-\frac{\frac{1}{\beta} + \bar{a}'}{\bar{a} - \bar{a}'}, 1\right) \end{aligned}$$

and

$$\begin{aligned} \hat{a}_U &= \beta((p - \varepsilon)\bar{a} + (1 - p + \varepsilon)\bar{a}') \in (\bar{a}, 1) \\ \alpha &= \frac{\hat{a}_U - \bar{a}}{\hat{a}_U + 1} \in (0, 1), \\ \alpha' &= \frac{\hat{a}_U - \bar{a}'}{\hat{a}_U + 1} \in (0, 1). \end{aligned}$$

(note that there exist parameters values consistent with these conditions since  $-1 < \frac{1-\beta(p-\varepsilon)\bar{a}}{\beta(1-p+\varepsilon)} < 0$  and  $-\frac{\frac{1}{\beta}+\bar{a}'}{\bar{a}-\bar{a}'} < 1$ ).

**Proof.** Given that  $p = \frac{1}{1+\bar{a}} \left(1 + \frac{\bar{a}}{\beta}\right)$ ,  $\beta < -1$  and  $\bar{a} > -1/\beta$ , computations show  $\hat{a}_U < 1$  for  $\varepsilon$  small enough (note that  $\bar{a}(\beta-1)\bar{a}' < \bar{a}(1-\beta) < (1+\bar{a}) - (\beta+\bar{a})\bar{a}$ ) and  $\hat{a}_U > \bar{a}$  follows from  $\bar{a}' < \frac{1-\beta(p-\varepsilon)}{\beta(1-p+\varepsilon)}\bar{a}$ . Hence,  $\hat{a}_U$  is the optimal action for an "uninformed" agent. The condition on  $q$  implies that the optimal action of the "informed" is  $-1$ . Lastly,  $\alpha$  and  $\alpha'$  satisfy the equilibrium conditions. The conditions  $\hat{a}_U > \bar{a}$  and  $\bar{a}' < 0$  imply  $\alpha, \alpha' \in (0, 1)$ . ■

The last case ( $\beta < 0$  and  $\bar{a} \leq 0$ ) is analogous to the above case  $\beta < 0$  and  $\bar{a} \geq 0$ , using  $+1$  instead of  $-1$  as a 2nd outcome when needed (to check this point, one just has to multiply every equilibrium condition by  $-1$ ). ■

### Proof of Proposition 5

We first give some notation. For every  $\bar{a}$  in  $[-1, 1]$ , the best response  $BR_{0,p}(\bar{a})$  is the solution of:

$$\max_a pu(a, 0) + (1-p)u(a, \bar{a}).$$

With the notation of the previous section, an element  $\tau$  in  $T$  is such that  $\tau_u$  is a Dirac measure on  $u$ . Then,  $\tau$  is characterized by a distribution on  $A$  (that is  $\tau_a$ ). With a slight abuse of notation, we identify an element  $\tau$  in  $T$  with its marginal  $\tau_a$  on  $A$ .

We check the following simple lemma:

**Lemma 2.** Consider an interval of actions  $[a_-, a_+]$  ( $0 \in [a_-, a_+]$ ), an action that is a best response to some beliefs on  $[a_-, a_+]$  putting at least probability  $p$  on 0 is an action in the interval  $[a'_-, a'_+]$  where

$$a'_- = \inf_{\bar{a} \in [a_-, a_+]} BR_{0,p}(\bar{a}) \text{ and } a'_+ = \sup_{\bar{a} \in [a_-, a_+]} BR_{0,p}(\bar{a}),$$

**Proof of the lemma.** The best response  $a$  of a player to belief on  $[a_-, a_+]$  putting at least probability  $p$  on 0 solves a FOC

$$pu'_a(a, 0) + (1-p) \int u'_a(a, \bar{a}) dP(\bar{a}) = 0,$$

where  $dP$  is some Borel measure on  $[a_-, a_+]$ . Notice that the LHS of this FOC is an integral over the family of functions  $pu'_a(a, 0) + (1-p)u'_a(a, \bar{a})$  (indexed by  $\bar{a}$ ). Furthermore,  $BR_{0,p}(\bar{a})$  is characterized as the solution of

$$pu'_a(a, 0) + (1-p)u'_a(a, \bar{a}) = 0.$$

The lemma follows. ■

We are now in a position to define the sequence of sets  $S_{0,p}^n$ . To this purpose, denote  $a_-^0 = -1$  and  $a_+^0 = +1$  and, for every  $n \geq 1$ , define iteratively the values  $a_-^n$  and  $a_+^n$  in  $[-1, 1]$  by

$$\forall n \geq 1, a_-^n = \inf_{\bar{a} \in [a_-^{n-1}, a_+^{n-1}]} BR_{0,p}(\bar{a}) \text{ and } a_+^n = \sup_{\bar{a} \in [a_-^{n-1}, a_+^{n-1}]} BR_{0,p}(\bar{a}),$$

(clearly,  $0 \in [a_-^n, a_+^n]$  and  $[a_-^n, a_+^n] \subset [a_-^{n-1}, a_+^{n-1}]$  for every  $n$ ).

- $T_p$  (that is  $S_{0,p}^0$ ) is the set of distributions on  $[a_-^0, a_+^0]$  putting at least probability  $p$  on 0
- An action that is a best response to some beliefs in  $S_{0,p}^0$  is an action in  $[a_-^1, a_+^1]$  (from the Lemma above)
- As every player is rational and has beliefs in  $S_{0,p}^0$ , the aggregate action is in  $[a_-^1, a_+^1]$ . Hence,  $\phi(S_{0,p}^0)$  is the set of distributions on  $[a_-^1, a_+^1]$ .
- $S_{0,p}^1 = \phi(S_{0,p}^0) \cap S_{0,p}^0$  is the set of distributions on  $[a_-^1, a_+^1]$  putting at least probability  $p$  on 0.

A comment about this argument: the key point here is that  $p$  has 2 effects on the transition between  $S_{0,p}^{n-1}$  and  $S_{0,p}^n$ : the "straight" effect that  $S_{0,p}^n$  is a set of distributions on a subset  $X$  of actions putting at least probability  $p$  on one specific action (the equilibrium), and the other effect (on which the iterative contraction argument relies), that the support  $X$  on the distributions in  $S_{0,p}^n$  shrinks with  $p$  ( $X$  decreases in  $p$ , for a given size of the support of  $S_{0,p}^{n-1}$ ).

We now iterate the argument:

- If  $S_{0,p}^{n-1}$  is the set of distributions on  $[a_-^{n-1}, a_+^{n-1}]$  putting at least probability  $p$  on 0, then an action that is a best response to some beliefs in  $S_{0,p}^{n-1}$  is an action in  $[a_-^n, a_+^n]$  (from the Lemma above)
- $\phi(S_{0,p}^{n-1})$  is then the set of distributions on  $[a_-^n, a_+^n]$
- $S_{0,p}^n = \phi(S_{0,p}^{n-1}) \cap S_{0,p}^{n-1}$  is the set of distributions on  $[a_-^n, a_+^n]$  putting at least probability  $p$  on 0.

By a standard argument, the two sequences  $a_-^n$  and  $a_+^n$  converge, and  $S_{0,p}^\infty$  is the set of distributions on  $[a_-^\infty, a_+^\infty]$  putting at least probability  $p$  on 0.  $S_{0,p}^\infty$  reduces to the equilibrium iff  $a_-^\infty = a_+^\infty = 0$ . If  $S_{0,p}^\infty$  does not reduce to the equilibrium, then every distribution on  $[a_-^\infty, a_+^\infty]$  putting at least probability  $p$  on 0 is a  $p$ -consensus distribution.

A necessary condition for  $a_-^\infty = a_+^\infty = 0$  is that  $BR_{0,p}$  is locally contracting at 0, that is:

$$|BR'_{0,p}(0)| < 1.$$

By the implicit functions theorem, we have:

$$BR'_{0,p}(\bar{a}) = -\frac{(1-p)u''_{a\bar{a}}(BR_{0,p}(\bar{a}), \bar{a})}{pu''_{aa}(BR_{0,p}(\bar{a}), 0) + (1-p)u''_{aa}(BR_{0,p}(\bar{a}), \bar{a})}.$$

Then (given  $BR_{0,p}(0) = 0$ )

$$BR'_{0,p}(0) = -\frac{(1-p)u''_{a\bar{a}}(0, 0)}{u''_{aa}(0, 0)},$$



and the condition  $|BR'_{0,p}(0)| < 1$  writes:

$$p > 1 - \left| \frac{u''_{aa}(0,0)}{u''_{a\bar{a}}(0,0)} \right|,$$

or, equivalently (differentiating the FOC  $u'_a(BR(\bar{a}), \bar{a}) = 0$  at  $(0,0)$ ):

$$p > 1 - \frac{1}{|BR'(0)|}. \quad (6)$$

On the other hand, a sufficient condition for  $a_-^\infty = a_+^\infty = 0$  is that  $BR_{0,p}$  is globally contracting:

$$\forall \bar{a} \in [-1, 1], |BR'_{0,p}(\bar{a})| < 1.$$

If  $M < 1$ , then this condition holds true for every value of  $p$  and  $\hat{p} = 0$ . We assume  $M \geq 1$  from now on. We have (given  $u''_{aa} < 0$ ):

$$|BR'_{0,p}(\bar{a})| = \frac{(1-p) |u''_{a\bar{a}}(BR_{0,p}(\bar{a}), \bar{a})|}{p |u''_{aa}(BR_{0,p}(\bar{a}), 0)| + (1-p) |u''_{aa}(BR_{0,p}(\bar{a}), \bar{a})|},$$

and  $|BR'_{0,p}(\bar{a})| < 1$  writes:

$$\frac{p}{1-p} > \left( \left| \frac{u''_{a\bar{a}}(BR_{0,p}(\bar{a}), \bar{a})}{u''_{aa}(BR_{0,p}(\bar{a}), \bar{a})} \right| - 1 \right) \left| \frac{u''_{aa}(BR_{0,p}(\bar{a}), \bar{a})}{u''_{aa}(BR_{0,p}(\bar{a}), 0)} \right|.$$

The RHS of this inequality is smaller than  $(M-1)m \geq 0$ . It follows that the sufficient condition for convergence holds for every  $p$  such that  $\frac{p}{1-p}$  is above this upper bound. Hence, the degree  $\hat{p}$  of  $p$ -stability satisfies:

$$\frac{\hat{p}}{1-\hat{p}} < (M-1)m,$$

or, equivalently:

$$\hat{p} < 1 - \frac{1}{1+(M-1)m}. \quad (7)$$

The existence of  $\hat{p}$  is shown in Proposition 2. Inequalities (6) and (7) imply the first part of the proposition. Propositions 1 and 3 imply the result on  $p$ -consensus distributions. ■

### Proof of Proposition 8.

As the economy is sequentially regular, using the implicit function theorem, we obtain the existence of a neighborhood  $N(\hat{\pi}_2, \varepsilon)$  of  $\hat{\pi}_2$  such that evaluated at the PFE  $(\hat{\pi}_1, \hat{q}, \hat{\pi}_2)$ ,

$$d\pi_2 = - \left( \int \partial_{\pi_2} \hat{\mathbf{x}}_2^i(\hat{\mathbf{x}}_1, \hat{\mathbf{y}}) di \right)^{-1} \int [\partial_y \hat{\mathbf{x}}_2^i d\hat{\mathbf{y}}^i + \partial_{x_1} \hat{\mathbf{x}}_2^i d\hat{\mathbf{x}}_1^i] di$$

where  $d\hat{\mathbf{y}}^i = \partial_{\pi_1} \hat{\mathbf{y}}^i d\pi_1 + \partial_q \hat{\mathbf{y}}^i dq + \partial_{\pi_2} \hat{\mathbf{y}}^i d\mathbf{f}^i$  and  $d\hat{\mathbf{x}}_1^i = \partial_{\pi_1} \hat{\mathbf{x}}_1^i d\pi_1 + \partial_q \hat{\mathbf{x}}_1^i dq + \partial_{\pi_2} \hat{\mathbf{x}}_1^i d\mathbf{f}^i$  for all  $i \in I$  and  $d\pi_1 = (\pi_1 - \hat{p}_1)$ ,  $dq = (q - \hat{q})$  and  $d\mathbf{f}^i = (\mathbf{f}^i - \hat{\pi}_2)$ .

As the PFE is sequentially regular,  $J_{11}^{-1} = \begin{pmatrix} z_1 & z_2 \\ z_3 & z_4 \end{pmatrix}$  exists. From the market clearing conditions in the spot markets at  $t = 1$ , by computation, it is checked that  $d\pi_1 = -z_1(\int \partial_{\pi_2} \hat{\mathbf{x}}_1^i d\mathbf{f}^i di) - z_2(\int \partial_{\pi_2} \hat{\mathbf{y}}^i d\mathbf{f}^i di) - z_3(\int \partial_{\pi_2} \hat{\mathbf{x}}_2^i d\mathbf{f}^i di) - z_4(\int \partial_{\pi_2} \hat{\mathbf{y}}^i d\mathbf{f}^i di)$ . It follows that setting

$$\mathbf{M}^i = \left( \int \partial_{\pi_2} \hat{\mathbf{x}}_2^i di \right)^{-1} \begin{bmatrix} (\int (\partial_y \mathbf{x}_2^i \partial_{p_1} \hat{\mathbf{y}}^i + \partial_{x_1} \hat{\mathbf{x}}_2^i \partial_{\pi_1} \hat{\mathbf{x}}_1^i) di) (z_1 \partial_{\pi_2} \hat{\mathbf{x}}_1^i + z_2 \partial_{\pi_2} \hat{\mathbf{y}}^i) \\ + (\int (\partial_y \mathbf{x}_2^i \partial_q \hat{\mathbf{y}}^i + \partial_{x_1} \hat{\mathbf{x}}_2^i \partial_q \hat{\mathbf{x}}_1^i) di) (z_3 \partial_{\pi_2} \hat{\mathbf{x}}_1^i + z_4 \partial_{\pi_2} \hat{\mathbf{y}}^i) \\ - (\partial_y \mathbf{x}_2^i \partial_{\pi_2} \hat{\mathbf{y}}^i + \partial_{x_1} \hat{\mathbf{x}}_2^i \partial_{\pi_2} \hat{\mathbf{x}}_1^i) \end{bmatrix}$$

for each  $i \in I$  yields the desired conclusion.

Let  $B(\bar{\varepsilon}) = \{x \in \mathfrak{R}_{++}^{L_2-1} : \|x - \hat{\pi}_2\| < \bar{\varepsilon}\}$ . By choosing  $N(\hat{\pi}_2, \varepsilon) \subseteq B(\bar{\varepsilon})$  so that  $\Pi_2^0 \subset B(\bar{\varepsilon})$ , the first-order approximation used in part (i) applies to all assignments of expectations  $\mathbf{f} : I \rightarrow \Pi_2^0$ . It follows that

$$\Pi_{2,0}^n = \{\pi_2 \in N(\hat{\pi}_2, \varepsilon) : d\pi_2 = \int \mathbf{M}^i d\mathbf{f}^i di, \text{ for some } \mathbf{f} : I \rightarrow \Pi_{2,0}^{n-1}\} \cap \Pi_{2,0}^{n-1}$$

for  $n = 1, 2, 3, \dots$ . Let  $\tilde{\mathbf{v}} : I \rightarrow \Pi_2^0$  be an assignment. For each  $i \in I$ , define  $v^i \in S(\bar{\varepsilon})$  by  $\text{sign } v^i = \text{sign } \tilde{\mathbf{v}}_i^i$ . Suppose there exists  $\bar{\varepsilon} > 0$  such that  $\|\int \mathbf{M}^i d\mathbf{v}^i di\| < \bar{\varepsilon}$ , for all assignment of expectations  $\mathbf{v} : I \rightarrow S(\bar{\varepsilon})$ . Observe that the map  $\mathbf{v}' : I \rightarrow S(\bar{\varepsilon})$  is an assignment. As  $\|\cdot\|$  is a monotone vector norm,  $\|\int \mathbf{M}^i d\tilde{\mathbf{v}}^i di\| \leq \|\int \mathbf{M}^i d\mathbf{v}^i di\| < \bar{\varepsilon}$ . Let  $\gamma(1) = \sup_{\mathbf{v} : I \rightarrow \Pi_2^0} \frac{\|\int \mathbf{M}^i d\mathbf{v}^i di\|}{\bar{\varepsilon}}$ . Then,  $\gamma(1) < 1$  and  $\Pi_{2,0}^1 \cap \Pi_{2,0}^0 \subseteq B(\gamma(1)\bar{\varepsilon})$ . For  $n = 1, 2, \dots$  define  $\gamma(n) = \sup_{\mathbf{v} : I \rightarrow \Pi_{2,0}^{n-1}} \frac{\|\int \mathbf{M}^i d\mathbf{v}^i di\|}{\bar{\varepsilon}}$ . Observe that  $\Pi_{2,0}^n \cap \Pi_{2,0}^{n-1} \subseteq B(\gamma(n)\bar{\varepsilon})$ . Further, as  $\|\cdot\|$  is a monotone vector norm,  $1 > \gamma(n-1) > \gamma(n)$ . It follows that  $\cap_{n \geq 0} \Pi_{2,0}^n \subseteq \cap_{n \geq 0} B(\gamma(n)\bar{\varepsilon}) \subseteq B(\gamma(0)^n \bar{\varepsilon}) = \{\hat{\pi}_2\}$ . As  $\{\hat{\pi}_2\} \subseteq \tilde{\Pi}_{2,0}$ , it follows that  $\tilde{\Pi}_{2,0} = \{\hat{\pi}_2\}$ .

Next, suppose there exists  $\bar{\varepsilon} > 0$  such that  $\|\int \mathbf{M}^i d\mathbf{v}^i di\| > \varepsilon$ , for all assignment of expectations  $\mathbf{v} : I \rightarrow S(\varepsilon)$  where  $\varepsilon \leq \bar{\varepsilon}$ . Then,  $\gamma(1) > 1$  and  $B(\gamma(1)\bar{\varepsilon}) \subseteq \Pi_{2,0}^n \cap \Pi_{2,0}^{n-1}$  so that  $B(\gamma(1)\bar{\varepsilon}) \subseteq \cap_{n \geq 0} \Pi_{2,0}^n$  so that  $\tilde{\Pi}_{2,0} \neq \{\hat{\pi}_2\}$ . Suppose there exists  $\bar{\varepsilon} > 0$  such that  $\|\int \mathbf{M}^i d\mathbf{v}^i di\| > \varepsilon$ , for all assignment of expectations  $\mathbf{v} : I \rightarrow S(\varepsilon)$ , for each  $\varepsilon \leq \bar{\varepsilon}$ . For  $\pi_2 \in N(\hat{\pi}_2, \bar{\varepsilon})$ , there exists  $\varepsilon' > 0$  and an assignment of expectations  $\mathbf{v} : I \rightarrow N(\hat{\pi}_2, \varepsilon')$  such that  $\int \mathbf{M}^i d\mathbf{v}^i di = \pi_2$ . Denote the corresponding set  $M_{\hat{\pi}_2, \mathbf{v}}$  and let  $M_{\hat{\pi}_2} = \cup_{\mathbf{v} \in \mathbf{V}} M_{\hat{\pi}_2, \mathbf{v}}$ . Note that  $M_{\hat{\pi}_2} \subseteq N(\hat{\pi}_2, \varepsilon')$ . Hence, for any  $\pi_2$  in  $N(\hat{\pi}_2, \varepsilon)$ , whenever  $\varepsilon$  is small enough,  $M_{\hat{\pi}_2} \subseteq N(\pi_2, \varepsilon')$ , so that  $M_{\hat{\pi}_2} = M_{\pi_2}$ ; moreover,  $M_{\pi_2} \subseteq \tilde{\Pi}_{2,0}$  for each  $\pi_2 \in N(\hat{\pi}_2, \varepsilon')$ , and hence,  $N(\hat{\pi}_2, \varepsilon'') \subseteq \tilde{\Pi}_{2,0}$  for all  $\varepsilon'' \leq \bar{\varepsilon}$ . By assumption,  $\tilde{\Pi}_{2,0} \subseteq N(\hat{\pi}_2, \bar{\varepsilon})$ , by continuity in  $p$ , there exists  $\tilde{p} > 0$  such that for all  $p < \tilde{p}$ ,  $\tilde{\Pi}_{2,p} \subseteq N(\hat{\pi}_2, \bar{\varepsilon})$ .

To check that the condition that  $\|\int \mathbf{M}^i d\mathbf{v}^i di\| < \bar{\varepsilon}$  is invariant to the choice of the second period numeraire, note that multiplying all prices by the same positive scalar  $\beta > 0$  implies that  $\bar{\varepsilon}$  on the right hand side of the condition is now  $\beta \bar{\varepsilon}$  while the expression on the left-hand side is equal to

$$\left\| \int -\beta \left( \int \partial_{\pi_2} \hat{\mathbf{x}}_2^i (\hat{\mathbf{x}}_1, \hat{\mathbf{y}}) di \right)^{-1} [\partial_y \hat{\mathbf{x}}_2^i d\hat{\mathbf{y}}^i + \partial_{x_1} \hat{\mathbf{x}}_2^i d\hat{\mathbf{x}}_1^i] di \right\| = \beta \left\| \int \mathbf{M}^i \mathbf{v}^i di \right\|$$

for all  $\mathbf{v} : I \rightarrow S(\bar{\varepsilon})$  which implies that the condition itself remains unchanged. ■

### Proof of Proposition 9.

For a fixed assignment of expectations  $\mathbf{v} : I \rightarrow \Pi_2^0$ ,  $\Pi_2^0 = N(\hat{\pi}_2, \varepsilon)$ , for some sequentially regular PFE  $(\hat{\pi}_1, \hat{q}, \hat{\pi}_2)$ ,

$$\left\| \int \mathbf{M}^i \mathbf{v}^i di \right\| = \left\| - \left( \int \partial_{\pi_2} \hat{\mathbf{x}}_2^i(\hat{\mathbf{x}}_1, \hat{\mathbf{y}}) di \right)^{-1} \int [\partial_y \hat{\mathbf{x}}_2^i d\hat{\mathbf{y}}^i + \partial_{x_1} \hat{\mathbf{x}}_2^i d\hat{\mathbf{x}}_1^i] di \right\|.$$

As market clearing in both periods is common knowledge,  $\int \hat{\mathbf{y}}^i di = 0$  and  $\int \hat{\mathbf{x}}_1^i di = \bar{\mathbf{w}}_1$  and therefore,  $\int d\hat{\mathbf{y}}^i di = \int d\hat{\mathbf{x}}_1^i di = 0$ . If both

$$\partial_y \hat{\mathbf{x}}_2^i(\hat{\mathbf{x}}_1^i, \hat{\mathbf{y}}^i(\hat{\pi}_1, \hat{q}, \hat{\pi}_2)) = \partial_y \hat{\mathbf{x}}_2^j(\hat{\mathbf{x}}_1^j, \hat{\mathbf{y}}^j(\hat{\pi}_1, \hat{q}, \hat{\pi}_2))$$

and

$$\partial_{x_1} \hat{\mathbf{x}}_2^i(\hat{\mathbf{x}}_1^i, \hat{\pi}_2, \hat{\mathbf{y}}^i) = \partial_{x_1} \hat{\mathbf{x}}_2^j(\hat{\mathbf{x}}_1^j, \hat{\pi}_2, \hat{\mathbf{y}}^j)$$

for all  $i, j \in I$ ,  $\int [\partial_y \hat{\mathbf{x}}_2^i d\hat{\mathbf{y}}^i + \partial_{x_1} \hat{\mathbf{x}}_2^i d\hat{\mathbf{x}}_1^i] di = 0$ , which, in turn, implies that  $\left\| \int M^i \mathbf{v}^i di \right\| = 0$  for every  $\mathbf{v} : I \rightarrow S(\varepsilon)$ , and every  $\varepsilon > 0$  such that  $S(\varepsilon) \subset \Pi_2^0$ . Therefore, for  $n = 1, 2, \dots$ ,  $\Pi_2^n = \{\hat{\pi}_2\}$ . By continuity of  $\|\cdot\|$ , there is an  $\tilde{\varepsilon}_1 > 0$  such that if

$$\max \left\{ \left\| \partial_y \hat{\mathbf{x}}_2^i(\hat{\mathbf{x}}_1^i, \hat{\pi}_2, \hat{\mathbf{y}}^i) - \partial_y \hat{\mathbf{x}}_2^j(\hat{\mathbf{x}}_1^j, \hat{\pi}_2, \hat{\mathbf{y}}^j) \right\|, \left\| \partial_{x_1} \hat{\mathbf{x}}_2^i(\hat{\mathbf{x}}_1^i, \hat{\pi}_2, \hat{\mathbf{y}}^i) - \partial_{x_1} \hat{\mathbf{x}}_2^j(\hat{\mathbf{x}}_1^j, \hat{\pi}_2, \hat{\mathbf{y}}^j) \right\| \right\} < \tilde{\varepsilon}_1$$

for all  $i, j \in I$ , there exists  $\varepsilon > 0$  such that  $S(\varepsilon) \subset \Pi_2^0$  and  $\left\| \int M^i \mathbf{v}^i di \right\| < \varepsilon$ , for all  $\mathbf{v} : I \rightarrow S(\varepsilon)$  and by Proposition 2,  $\tilde{\Pi}_{2,0} = \tilde{\Pi}_{2,p} \{\hat{\pi}_2\}$ ,  $0 \leq p \leq 1$ . Next, note that using the expression for  $M^i$  derived in Proposition 1, by continuity of  $\|\cdot\|$ , there is an  $\tilde{\varepsilon}_2 > 0$  such that if  $\max \{ \|\partial_{\pi_2} \hat{\mathbf{y}}^i\|, \|\partial_{\pi_2} \hat{\mathbf{x}}_1^i\| \} < \tilde{\varepsilon}_2$ , for all  $i \in I$ , there exists  $\varepsilon > 0$  such that  $S(\varepsilon) \subset \Pi_2^0$  and  $\left\| \int M^i \mathbf{v}^i di \right\| < \varepsilon$ , for all  $\mathbf{v} : I \rightarrow S(\varepsilon)$ . Finally, the set  $\tilde{\varepsilon} = \min \{\tilde{\varepsilon}_1, \tilde{\varepsilon}_2\}$ . ■