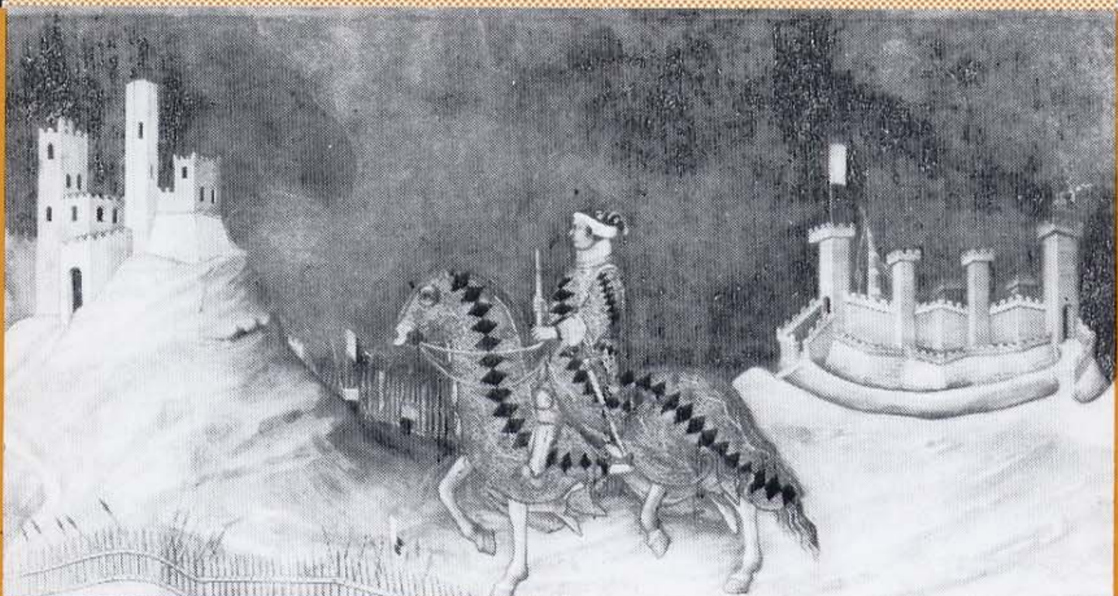
UNIVERSITÀ DEGLI STUDI DI SIENA

QUADERNI DEL DIPARTIMENTO DI ECONOMIA POLITICA

Leonardo Boncinelli

Global vs. Local Information

n.520 - Ottobre 2007



Abstract - In this paper I apply stochastic stability to compare local information to global information in terms of welfare. Under global information agents potentially imitate anyone else, while under local information choices are grouped into information sets and agents can observe and hence imitate only those within their own information set.
The welfare evaluation of information is ambiguous over finite time horizons, while in the long run less (more) information is better in the presence of pure negative (positive) spillovers. However, when a selection issue is considered a further ambiguity emerges making the comparison, in general, uncertain.

JEL classification: D01; D80; D81; D85; Z13
Keywords: imitation; local information; global information; stochastic stability

Leonardo Boncinelli, Dipartimento di Economia Politica, Università di Siena

1 Introduction

With the aim of illustrating the subject matter of this paper, consider the following example. There are two areas, and in each of them several investment opportunities are available. Every agent has to decide where to invest. The profitability of each investment opportunity depends (positively or negatively) upon how many people make that choice. Agents make their investment decisions by looking at other agents and imitating some of them, for instance those who gain higher (or the highest) returns. However, not all agents can be observed, but only those who invest in the same area. With a very small probability agents make mistakes and choose whatever investment in whatever area.

Suppose that there is a public authority which can intervene making the return on any realized investment a public information. Is such a policy welfare enhancing? Or instead is it better to leave the existing information structure? The present paper tries to answer these questions with a certain generality.

The question whether public information is socially beneficial has been addressed in the literature since long time (Hirshleifer, 1971). Recently, Morris and Shin (2002) show that in a model with complementarities public information can be detrimental, since agents are very reactive to it and this can magnify the damage done by any noise. In the investment game of Angeletos and Pavan (2004) the frequent provision of public information is instead beneficial, even if that may lead to an increase in volatility. Angeletos and Pavan (2007) relate the ambiguous role played by more precise public information to discrepancies between the effects on the equilibrium use of information and the effects on the efficient use of information.

This paper faces a similar question from a radically different perspective. Here there are no noisy signals representing private and public information. Each agent has certain but incomplete information coming from her personal experience, that is her current choice. In this setting, complete (or global, as generally denominated here) information can be interpreted as the result of the disclosure of the whole of information by a public authority. Furthermore, the model I build is not based on standard game theory but is instead of evolutionary type, with individual behavior driven by the imitation of successful agents. In some cases stochastic stability is applied to get a unique prediction for the long run behavior of the system, as widely used in economics since Young (1993) and Kandori et al. (1993) (see Young (1998) for a comprehensive treatment). Probability distributions over states relative to different information structures are then compared in order to rank them, if possible, in terms of welfare.

The main issue in the paper is how information is used to solve a *coordination issue*. Positive (negative) spillovers occur when the more people choose an alternative, the higher (the lower) the individual payoff; furthermore, spillovers are denominated pure if any two alternatives chosen by the same number of agents are worth the same. It is self-evident that in the presence of pure positive spillovers coordination is welfare enhancing, and the converse is true in the presence of negative spillovers. Since agents choose only what they are informed about, having the same information is intuitively better when they have to coordinate, while having different information is intuitively better when they have to miscoordinate. Indeed, more information makes more agents behave according to the same rule, but it can also change such rule from deterministic to probabilistic or, more in general, it can disperse the mass of probability. This generates mixing effects which counteract the previous intuition.

However, when working with the stochastically stable distributions these mixing effects disappear since the recurrent states are homogeneous within each information set. This allows to establish the welfare enhancing effect of more (less) information in the presence of pure positive (negative) spillovers.

The final part of the paper deals with a more general class of interactions. When preferences do not depend solely upon how many do something but also upon what they choose, a *selection issue* emerges and the previous result ceases to hold. The reasons of such a failure are exemplified and, then, a general condition under which global information is never welfare superior to local information is identified.

The next sections are an attempt to formalize these sketched issues and to analize their interplay in shaping the results. More precisely, section 2 presents the model and introduces preliminaries. Section 3 compares different information structures in the presence of pure negative spillovers, while section 4 briefly does the same in the presence of pure positive spillovers. Section 5 explores a more general class of interactions where a selection issue emerges. Section 6 concludes.

2 The model

Consider the quadruple $< \mathcal{M}, \mathcal{C}, \mathcal{I}, \succcurlyeq >$, where $\mathcal{M}$ is a finite set of agents of cardinality m, $\mathcal{C}$ is a finite set of choices (or alternatives) of cardinality n, $\mathcal{I} = \{I_1, I_2, \ldots, I_l\}$ is a partition of $\mathcal{C}$ in information sets, and $\succcurlyeq$ is a preference relation over $\mathcal{C} \times \{1, 2, \ldots, m\}$, with $(a, k) \succcurlyeq (b, h)$ to be interpreted as "choosing alternative a when k agents choose alternative a is at least as good as choosing alternative b when h agents choose alternative b". The relations $\succ$ and $\sim$ are derived as usual.

Define *global information* as the partition $\mathcal{I}_g$ with a unique information set comprising all alternatives. An information structure[1] $\mathcal{I}_1$ is said to be

[1]Please note that information structures in this framework have nothing to do with information structures in game theory, as originating from the seminal work of Aumann (1976).

more detailed than an information structure $\mathcal{I}_2$ when $\mathcal{I}_2$ is a strictly finer partition than $\mathcal{I}_1$.

The preference relation is said to exhibit *negative spillovers* if for any $a \in \mathcal{C}$, for any $1 \leq k < m$, $(a, k) \succ (a, k+1)$. Analogously, the preference relation is said to exhibit *positive spillovers* if for any $a \in \mathcal{C}$, for any $1 \leq k < m$, $(a, k+1) \succ (a, k)$.[2] The preference relation is said to exhibit *pure negative spillovers* (*pure positive spillovers*) if it exhibits negative spillovers (positive spillovers) and $(a, k) \sim (b, k)$ for any a, $b \in \mathcal{C}$ and for any $1 \leq k \leq m$.

A state of the system is defined as a function $x : \mathcal{M} \longrightarrow \mathcal{C}$ assigning a choice to every agent.

Given a state x, an agent r *observes* any agent s such that $x(r)$ and $x(s)$ belong to the same information set. Call $O_r(x)$ the set of agents observed by r in state x and, with some abuse of notation, let $x[O_r(x)]$ be the set of their choices. By definition $r \in O_r(x)$ and hence $x(r) \in x(O_r(x))$ for any x.

Consider the following class of imitation rules.

Imitation rule. The behavioral rule of agents satisfies the following two properties:

(P1) if an alternative is chosen by agent r with positive probability at the next time then it belongs to $x(O_r(x))$;

(P2) if an alternative is maximally preferred within $x(O_r(x))$ then it is chosen with positive probability by r.[3]

Note that non-maximally preferred alternatives in $x(O_r(x))$ can be chosen with zero probability. This class encompasses virtually all forms of imitation and mixtures between imitation[4] and inertia.[5] As particular cases there are most of the imitation rules used in economics, such as the imitation of those who performed the best (Vega-Redondo, 1997) the imitation of those who performed better with a probability increasing in the payoff differential (Schlag, 1998, 1999) and the imitation of anyone with a probability increasing in the observed payoff (Schlag, 1999). Every agent in $\mathcal{M}$ is

[2]These definitions are adapted from Cooper and John (1988).

[3]This last condition allows to avoid cycles of imitative behaviors.

[4]One can think about imitation as naturally leading to copy out the observed best alternative. However, agents might be able only to compare their own choices with an observed choice, but not observed choices between them. In such a case it would not be possible to determine the best, and agents might be reasonably supposed to copy any better choice with positive probability. The last imitation rule is also plausible when an agent can potentially observe any agent in her own information set, and she is randomly matched with only one of them, choosing to copy her choice if not inferior. As an extreme case, agents might be unable even to compare observed choices with their own, and hence they might imitate anyone they observe with positive probability.

[5]Inertia can be included by simply assuming that agents always keep on making their own choice with a certain probability. Note that pure imitation is admitted but pure inertia not.

assumed to follow the same behavioral rule in this class. Actually, any result of the paper holds even under heterogeneous behavioral rules. However, it is clear that more information can hinder coordination if people use the additional information differently. What I want to point out is that even if people adopt the same behavioral rule, still another kind of hindrance exists due to the possible increase in the dispersion of choices.

The dynamic process so defined is a Markov chain, call it M^0. Define X^* as the set of states which are homogeneous within each information set, formally $X^* = \{x \in X : \forall r, s \in \mathcal{M}, (\exists I_i \in \mathcal{I} : x(r) \in I_i \wedge x(s) \in I_i) \Rightarrow x(r) = x(s)\}$. The following lemma is easily established.

Lemma 1. *The recurrent classes of the Markov chain M^0 are the singletons containing the states in X^*.*

Proof. Clearly the states in X^* are absorbing states of M^0, since any agent observes nothing different to imitate.

Take a state out of X^*. Choose in each information set a homogeneous group of agents who are choosing a maximally preferred alternative therein. Each of those agents has a positive probability to keep on making the same choice, by (P2). Any other agent imitates one of them with positive probability, again by (P2). Since the number of agents is finite, from any state there is a positive probability to reach in a single period a state in X^*. $\square$

The initial condition clearly matters for establishing the final state. For instance, if the system starts from whatever absorbing state, it will remain there forever. In other words, the system is not ergodic. In order to get a unique prediction and to add a layer of realism, perturbations of individual choices are introduced. In particular, at any time each agent with probability $\epsilon > 0$ makes a random choice, with any alternative in $\mathcal{C}$ equally probable. This perturbed dynamic process, which will be denoted by M^ϵ, is an irreducible and aperiodic Markov chain.[6] Let $\mu^{t,\epsilon}(x|x^0)$ be the probability to be in state x at time t given x^0 as initial state. By known results, as time goes to infinity the system converges to the unique stationary distribution, denoted by $\mu^{*,\epsilon}$, irrespectively of the initial condition. More precisely, if X is the state space, i.e. the set containing any conceivable function $x : \mathcal{M} \longrightarrow \mathcal{C}$, then μ^ϵ is a probability distribution over X, where $\mu^\epsilon(x)$ represents, as time goes to infinity, both the proportion of time that the system will have spent on x and the probability that the system will be exactly at x.

However, being the stationary distribution hard to compute in general, it has become standard to consider its limit for a vanishing ϵ. By so doing the stochastically stable distribution is obtained, denoted by μ^*, and

[6] A Markov chain is said irreducible when there is a positive probability of moving from any state to any other state in a finite number of periods, and it is said aperiodic when for every state x unity is the greatest common divisor of the set of all the integers r such that there is a positive probability of moving from x to x in exactly r periods.

some techniques have been developed to determine which states are visited with positive probability in such distribution (Young, 1993; Kandori et al., 1993; Ellison, 2000). Hereafter, I will use the pedices 1, 2 and g for probability distributions to refer them to the information structures $\mathcal{I}_1$, $\mathcal{I}_2$, $\mathcal{I}_g$ respectively.

The following lemma provides a characterization of the stochastically stable states of any information structure different from global information, independently of the preference relation.

Lemma 2. *If the information structure is different from global information, then the stochastically stable states of the perturbed Markov chain M^ϵ are the states in X^*.*

Proof. By lemma 1 it is known that a state out of X^* cannot be stochastically stable, since it is neither an absorbing state.

At least one state in X^* must be stochastichally stable. Call it x^*. It is known (see, for instance, Lemma 1 in Ben-Shoham et al. (2004)) that if the edge (x_1, x_2) has resistance 1,[7] x_1 is stochastically stable and x_2 is an absorbing state, then x_2 is stochastically stable. Therefore, it is sufficient to prove that there is a path of absorbing states from x^* to any other state x' in X^* with the resistance of any edge equal to 1. Let d be the distance between x^* and x', that is the number of agents who make in x^* a different choice from x'. I will show how to find an appropriate path of absorbing states which reduces the distance to x' by 1. Then the result follows by induction.

Suppose $x^*(r) = a \neq x'(r) = b$, with $a \in I_i$ and $b \in I_j$. Case i): if no agent in x^* makes a choice in I_j different from b, then a single perturbation can make agent r choose b, so reaching an absorbing state with only $d-1$ agents making a different choice with respect to x'. Case ii): if instead other agents in x^* make a choice in I_j different from b, first note that those agents cannot make the same choice in x', otherwise x' would not be homogeneous within each information set. Second, note that there exists an information set different from I_j since the information structure is different from global information. Then each of those agents can move out of I_j with a single perturbation, every time reaching an absorbing state with no more than d agents making a different choice with respect to x'. At most, when all of them have left, the same reasoning as in case i) applies. $\square$

In the following sections, welfare evaluations are made comparing distributions over states. In particular, a distribution μ_1 is said *welfare superior* to a distribution μ_2 when for every agent μ_1 first-order stochastically dominates μ_2.

[7]The *resistance* from state x_1 to state x_2 is defined as the minimum number of perturbations required to move from x_1 to x_2 with positive probability.

3 Pure negative spillovers

In the presence of pure negative spillovers, coordination on the same choice is harmful. It seems intuitive that the more detailed the information structure, the more agents observe each other and coordinate through imitation on the same choice. The following example illustrates.

Example 1. Suppose $\mathcal{C} = \{a, b\}$, $m = 2$ and the preference relation exhibits pure negative spillovers. Furthermore, assume an imitation rule which in case of ties provides to choose between the alternatives with equal probability. Define $\mathcal{I}_1 = \{\{a\}, \{b\}\}$.

If both agents are making the same choice in the initial state x^0, then $\mu_g^{1,\epsilon}(\cdot|x^0) = \mu_1^{1,\epsilon}(\cdot|x^0)$. Let α be the probability they miscoordinate under such unique distribution. If instead the agents are making different choices in the initial state, then the probability β_1 to miscoordinate under $\mathcal{I}_1$ is larger than the probability β_g to miscoordinate under $\mathcal{I}_g$ (since $\beta_1 = (1-\epsilon/2)^2 + (\epsilon/2)^2 > 1/2 = \beta_g$). Call p_1^t and p_g^t the probability to miscoordinate after t periods under, respectively, $\mathcal{I}_1$ and $\mathcal{I}_g$. These probabilities can be defined recursively: $p_1^{t+1} = p_1^t\beta_1 + (1-p_1^t)\alpha$ and $p_g^{t+1} = p_g^t\beta_g + (1-p_g^t)\alpha$, with $p_1^1 \geq p_g^1$ since they are either both equal to α or $p_1^1 = \beta_1$ and $p_g^1 = \beta_g$. So, $p_1^t > p_g^t$ for any $t \geq 2$ and $\mu_1^{t,\epsilon}(\cdot|x^0)$ turns out to be welfare superior to $\mu_g^{t,\epsilon}(\cdot|x^0)$ for $t \geq 2$ whatever the initial state.

However, the class of imitation rules that is considered in this paper is large enough for the previous intuition to be proven incorrect. In particular, a more detailed information structure does not necessarily yield a welfare inferior distribution when the time horizon is finite. More detailed information makes more agents align their expected choices at the next time. This per se favors coordination. However, more information can also cause the behavioral rule to share out the probability of the next choice among more alternatives. In such a case, the increase in the dispersion of choices reduces coordination. Examples 2 and 3 provide some simple instances. Such examples are arranged so that you can use behavioral rules where only not inferior choices are copied with positive probability. This to be convinced that the source of problems would not be eliminated by the restriction to this kind of rules.

Example 2. Take $\mathcal{C} = \{a, b\}$, $m = 3$, a preference relation with pure negative spillovers and $\mathcal{I}_1 = \{\{a\}, \{b\}\}$. Suppose also that the imitation rule of every agent is such that she keeps on making her previous choice with some positive probability. To simplify calculations, let ϵ be so small to be negligible. Take as initial condition a state x^0 where two agents choose a and one agent chooses b. The probability that one of the agents initially choosing a finds herself alone under $\mu_g^{1,\epsilon}(\cdot|x^0)$ is positive, while the same probability is zero under $\mu_1^{1,\epsilon}(\cdot|x^0)$. For such an agent $\mu_1^{1,\epsilon}(\cdot|x^0)$ does not first-order stochastically dominate $\mu_g^{1,\epsilon}(\cdot|x^0)$.

Example 3. Consider $\mathcal{C} = \{a, b, c\}$, $m = 6$, a preference relation with pure negative spillovers and $\mathcal{I}_1 = \{\{a\}, \{b\}, \{c\}\}$. Assume an imitation rule without inertia where every observed choice strictly better than her own is imitated with positive probability. Take an initial state where an agent chooses a, two agents choose b and the remaining three agents choose c. Let ϵ be so small to be negligible. The probability that one of the agents initially choosing c finds herself alone is positive under $\mu_g^{1,\epsilon}(\cdot|x^0)$ and zero under $\mu_1^{1,\epsilon}(\cdot|x^0)$. Clearly $\mu_1^{1,\epsilon}(\cdot|x^0)$ does not first-order stochastically dominate $\mu_g^{1,\epsilon}(\cdot|x^0)$ for such an agent.

The previous examples raise some doubts about the choice of first-order stochastic dominance for welfare comparisons. Is it too demanding to be of any help for the evaluation of information? The following proposition shows this is not the case, at least when working with the stochastically stable distributions. Intuitively, the states where mixing effects are present occur with negligible probability in the very long run in the presence of rare perturbations. In fact, they are not recurrent since the only recurrent states are those which are homogeneous within each information set. Therefore those states become irrelevant and the welfare superiority of less detailed information structures can be established.

Proposition 1. *Suppose the preference relation exhibits pure negative spillovers. If $\mathcal{I}_1$ is more detailed than $\mathcal{I}_2$, then μ_2^* is welfare superior to μ_1^*.*

Proof. I consider $\mathcal{I}_1$ and $\mathcal{I}_2$ such that $\mathcal{I}_1 = \{I_1, I_2, \ldots, I_i, \ldots, I_l\}$, $\mathcal{I}_2 = \{I_1^{'}, I_1^{''}, I_2, \ldots, I_i, \ldots, I_l\}$ with $I_1 = I_1^{'} \cup I_1^{''}$. In fact, any comparison between more/less detailed information structures can be reduced to a sequence of such simpler comparisons. To simplify the following exposition, define $p(A, a|\mu)$ and $p(A, B|\mu)$ as the probability under distribution μ that all and only the agents in set $A \subseteq \mathcal{M}$ choose, respectively, alternative a and alternatives in $B \subseteq \mathcal{C}$; define also $p(A_1, a_1; A_2, a_2|\mu)$ as the probability under distribution μ that all and only the agents in A_1 choose alternative a_1 and all and only the agents in A_2 choose alternative a_2.

Take whatever initial state x^0 and whatever non-empty subset A of agents. Then $p(A, a|\mu_1^{1,\epsilon}) = p(A, a|\mu_2^{1,\epsilon})$ for any $a \in I_i \neq I_1$. This is so because any agent currently choosing out of I_i will make choice a with probability ϵ/n independently of the rest, while agents currently choosing in I_i will make choice a depending on the state of the system restricted to I_i (which is the same under the two distributions). Being this true for any subset A of players and for any $a \in I_i$, the probability distribution over the states of the system restricted to I_i at the next time is the same under $\mathcal{I}_1$ and $\mathcal{I}_2$. Therefore the same reasoning can be applied indefinitely obtaining that $p(A, a|\mu_1^{t,\epsilon}) = p(A, a|\mu_2^{t,\epsilon})$ for any $a \in I_i \neq I_1$ and for any t. As a consequence, this is still true in the limit for $t \to \infty$, i.e. $p(A, a|\mu_1^{*,\epsilon}) = p(A, a|\mu_2^{*,\epsilon})$, and in the limit for $\epsilon \to 0$, i.e. $p(A, a|\mu_1^*) = p(A, a|\mu_2^*)$.

Similarly, $p(A, I_1|\mu_1^{t,\epsilon}) = p(A, I_1|\mu_2^{t,\epsilon})$ for any t, and hence $p(A, I_1|\mu_1^*) = p(A, I_1|\mu_2^*)$ too. In fact, any agent choosing out of I_1 makes a choice in I_1 with probability $\epsilon||I_1||/n$ independently of the rest, while agents choosing in I_1 keep on choosing in I_1 with probability $1 - \epsilon(n - ||I_1||)/n$ independently of the rest.

Now, by lemma 1 the following two equalities hold, and by lemma 2 each term of the sum in the right-hand side of the second equality is positive:

$$\begin{aligned} p(A, I_1|\mu_1^*) &= \sum_{a \in I_1} p(A, a|\mu_1^*) \\ p(A, I_1|\mu_2^*) &= \sum_{\substack{A_1 \cup A_2 = A, \\ A_1 \cap A_2 = \emptyset}} \sum_{\substack{a_1 \in I_1', \\ a_2 \in I_1''}} p(A_1, a; A_2, b|\mu_2^*) \end{aligned}$$

Finally, take any agent r and consider every subset A_r of players containing r. Therefore:

$$\begin{aligned} p(A_r, a|\mu_1^*) &= p(A_r, a|\mu_2^*) \quad \text{if } a \notin I_1 \\ \sum_{a \in I_1} p(A_r, a|\mu_1^*) &= \sum_{\substack{A_1 \cup A_2 = A_r, \\ A_1 \cap A_2 = \emptyset}} \sum_{\substack{a_1 \in I_1', \\ a_2 \in I_1''}} p(A_1, a_1; A_2, a_2|\mu_2^*) \end{aligned}$$

Now, since by the assumption of pure negative spillovers agent r is better off in any state where the agents in A_r are actually split into alternatives a_1 and a_2 instead of being all grouped in a single alternative, first-order stochastic dominance follows. $\square$

The following proposition weakens the assumption on the preference relation, which is no more required to exhibit pure negative spillovers globally. In fact, what is now required is that when all the m agents make the same choice they get the least preferred outcome whatever the choice. Note that there is no requisite about the preference relation restricted to $\mathcal{C} \times \{1, 2, \ldots, m-1\}$. Hence, for example, a group of $m - 1$ agents can be better off than an isolated agent.

Proposition 2. *Suppose the preference relation is such that $(a, k) \succ (b, m) \sim (c, m)$ for any $a, b, c \in \mathcal{C}$ and for any $1 \leq k < m$. If $\mathcal{I}_1 \neq \mathcal{I}_g$, then μ_1^* is welfare superior to μ_g^*.*

Proof. Suppose $\mathcal{I}_1 = \{I_1, I_2, \ldots, I_l\}$. Let $\mathcal{A} = \{A_1, A_2, \ldots, A_l\}$ be a generic partition of set M in l subsets (some of them possibly empty). Define $p(A_1, a_1; A_2, a_2; \ldots; A_l, a_l|\mu)$ as the probability under distribution μ that all and only the agents in A_1 choose alternative a_1 and all and only the agents in A_2 choose alternative a_2 and ... and all and only the agents in A_l choose alternative a_l.

By lemma 1 the following two equalities hold, and by lemma 2 each term of the sum in the right-hand side of the second equality is positive:

$$\begin{aligned} 1 &= \sum_{a \in \mathcal{C}} p(M, a|\mu_g^*) \\ 1 &= \sum_{\mathcal{A}} \sum_{\substack{a_1 \in I_1, a_2 \in I_2, \\ \ldots, a_l \in I_l}} p(A_1, a_1; A_2, a_2; \ldots; A_l, a_l|\mu_1^*) \end{aligned}$$

Since by hypothesis any agent is better off when grouped with less than other $m-1$ agents, first-order stochastic dominance follows. $\square$

4 Pure positive spillovers

This section is specular to the previous one and therefore I will be much shorter in explanations. The same intuition as before applies to the coordination issue, with the only difference that in the presence of pure positive spillovers coordination is beneficial.

The following examples are simple modifications of those used for pure negative spillovers and illustrate analogous points. Again, you can use behavioral rules where only not inferior choices are copied with positive probability. Obviously, the conclusions are reversed but the same ambiguity about the evaluation of information over finite time horizons follows.

Example 1′. Consider example 1 after substituting pure negative spillovers with pure positive spillovers. All the reasoning remains exactly the same, with just the opposite result due to the different assumption on the preference relation.

Example 2′. Take $\mathcal{C} = \{a, b\}$, $m = 5$, a preference relation with pure positive spillovers and $\mathcal{I}_1 = \{\{a\}, \{b\}\}$. Suppose also that the imitation rule of every agent is such that she keeps on making her previous choice with some positive probability. To simplify calculations, let ϵ be so small to be negligible. Take as initial condition a state x^0 where two agents choose a and three agents choose b. The probability that one of the agents initially choosing a finds herself alone under $\mu_g^{1,\epsilon}(\cdot|x^0)$ is positive, while the same probability is zero under $\mu_1^{1,\epsilon}(\cdot|x^0)$. For such an agent $\mu_g^{1,\epsilon}(\cdot|x^0)$ does not first-order stochastically dominate $\mu_1^{1,\epsilon}(\cdot|x^0)$.

Example 3′. Consider $\mathcal{C} = \{a, b, c\}$, $m = 9$, a preference relation with pure positive spillovers and $\mathcal{I}_1 = \{\{a\}, \{b\}, \{c\}\}$. Assume an imitation rule without inertia where every observed choice strictly better than her own is imitated with positive probability. Take an initial state where four agents choose a, three agents choose b and the last two agents choose c. Let ϵ be so small to be negligible. The probability that one of the agents initially choosing c finds herself alone is positive under $\mu_g^{1,\epsilon}(\cdot|x^0)$ and zero under $\mu_1^{1,\epsilon}(\cdot|x^0)$. Clearly $\mu_g^{1,\epsilon}(\cdot|x^0)$ does not first-order stochastically dominate $\mu_1^{1,\epsilon}(\cdot|x^0)$ for such an agent.

Again, the stochastically stable distributions allow to rank information structures more easily. The following propositions $1'$ and $2'$ are the counterpart of propositions 1 and 2.

Proposition 1′. *Suppose the preference relation exhibits pure positive spillovers. If $\mathcal{I}_1$ is more detailed than $\mathcal{I}_2$, then μ_1^* is welfare superior to μ_2^*.*

Proof. The same as for proposition 1, with the reversed conclusion due to pure positive spillovers. □

Proposition 2′. *Suppose the preference relation is such that $(a, k) \prec (b, m) \sim (c, m)$ for any $a, b, c \in \mathcal{C}$ and for any $1 \leq k < m$. If $\mathcal{I}_1 \neq \mathcal{I}_g$, then μ_g^* is welfare superior to μ_1^*.*

Proof. The same as for proposition 2, with the reversed conclusion due to pure positive spillovers. □

5 Selection

Most of the analysis in the previous two sections has been carried out under the simplyfying assumption of a preference relation which depends uniquely upon the number of agents making a choice, not upon the choice itself. It is potentially interesting to investigate whether the previous results hold under weaker assumptions and, if it is not the case, what type of problems emerges. All the discussion in this section is developed considering only behavioral rules where inferior choices are not copied. This allows to interpret selection as standard in evolutionary theories: what has proven better tends to spread. Moreover, since this section mainly provides counterexamples, having restricted the class within which looking for strengthens the results. Finally, note that proposition 3 holds under general behavioral rules.

What can be said by assuming only negative spillovers instead of pure negative spillovers? Unfortunately, very little, as the next two examples show.

Example 4. Suppose there is an action in $\mathcal{C}$, call it a, which is very bad, so bad that $(a, 1) \prec (b, m)$ for any $b \in \mathcal{C}$. Suppose also that the preference relation exhibits negative spillovers. Assume an imitation rule where observed inferior choices are not copied. Consider global information. From the state where all agents choose a, call it $\tilde{x}$, a single perturbation is sufficient to move the system away, while from any other homogeneous state m perturbations are required to move the system to $\tilde{x}$. By a mutation counting argument it is intuitive that $\tilde{x}$ is not a stochastically stable state under global information. By lemma 2 instead, under any less detailed information structure each agent has a positive probability to choose a in the stochastically stable distribution. Therefore, μ_g^* is not welfare inferior to any other stochastically stable distribution.

One could argue that the crucial point in the previous example is that alternative a is dominated, and that therefore a single perturbation is not sufficient to move the system to $\tilde{x}$. One could hope that by requiring $(a, 1) \succ (b, m-1)$ the welfare superiority of less detailed information structures can be reasserted. Such a hope vanishes when considering the following example.

Example 5. Take $\mathcal{C} = \{a, b\}$, $m = 4$, a preference relation with negative spillovers and $\mathcal{I}_1 = \{\{a\}, \{b\}\}$. Suppose also that $(a, 2) \succ (b, 2)$, $(b, 1) \succ (a, 3)$ and that the least preferred case is $(b, 4)$, which happens in state $\tilde{x}$. Assume an imitation rule where observed inferior choices are not copied and which provides to keep on making the same choice with probability $(1 - \beta)$.

By symmetry, every state has the same probability in μ_1^*, therefore $\mu_1^*(\tilde{x}) = \frac{1}{16}$.

Under global information the stochastically stable states are $\tilde{x}$ and the state where every agent chooses a, call it $\hat{x}$. In fact, they are the only recurrent states by lemma 1, and a single perturbation is sufficient to move the system from one them to the other and viceversa. In the following I try to establish how large is their probability in the stochastically stable distribution.

It is known (Freidlin and Wentzell, 1984) that

$$\frac{\mu_g^{*,\epsilon}(\tilde{x})}{\mu_g^{*,\epsilon}(\hat{x})} = \frac{\displaystyle\sum_{F_{\tilde{x}} \in \mathcal{F}_{\tilde{x}}} L(F_{\tilde{x}})}{\displaystyle\sum_{F_{\hat{x}} \in \mathcal{F}_{\hat{x}}} L(F_{\hat{x}})}$$

where for a generic state x, F_x is an x-tree, that is a tree with vertex x and X as set of nodes, $\mathcal{F}_x$ is the set containing all x-trees, $L(F_x)$ is the likelihood of the x-tree F_x, that is $L(F_x) = \prod_{(x_1,x_2) \in F_x} T_{x_1 x_2}$ with $T_{x_1 x_2}$ the transition probability from x_1 to x_2. When ϵ tends to zero, the relevant trees for the previous ratio become those with the minimum number of perturbations.

The following figure represents the only type of $\hat{x}$-tree and the only type of $\tilde{x}$-tree which minimize the number of perturbations and, conditional to that, minimize the number of imitations of different behaviors. By $(k, m-k)$ in the figure I mean a state where k agents choose a and $m-k$ agents choose b. Beside each arrow there is the relevant term (without coefficient) expressing such transition probability.

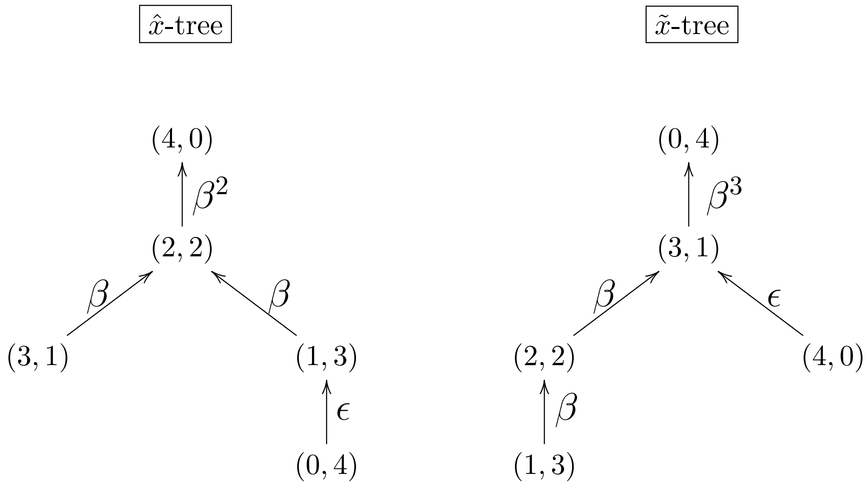


Figure 1: Stochastically stable distributions.

In $\sum_{F_{\tilde{x}} \in \mathcal{F}_{\tilde{x}}} L(F_{\tilde{x}})$ the term ϵ is multiplied by a polynomial in β whose lower term has degree 5. In $\sum_{F_{\hat{x}} \in \mathcal{F}_{\hat{x}}} L(F_{\hat{x}})$ the term ϵ is instead multiplied by a polynomial in β whose lower term has degree 4. Therefore, when inertia is sufficiently high, that is when β is sufficiently close to 0, $\mu_g^*(\tilde{x})/\mu_g^*(\hat{x})$ can be made as small as desired. In particular, when $\mu_g^*(\tilde{x})/\mu_g^*(\hat{x}) < 1/15$, $\mu_g^*(\tilde{x}) < 1/16$, so showing that μ_1^* is not welfare superior to μ_g^*.

The previous discussion concerned negative spillovers. What about positive spillovers? At first one could be tempted to assert that more information is even better from a welfare point of view in the presence of a selection issue, implicitly assuming that a larger amount of information helps selecting a better choice. Indeed, such a hasty conclusion would represent a serious mistake. Stochastic stability is able to select what is evolutionarily efficient, that is - to say that in biological terms - the best against invasions by mutants. It is well known (Young, 1993, 1998) that stochastic stability does not ensure efficiency. However, one could hope that it is at least able to select a better choice under global information than under less detailed information structures, which are substantially based on some kind of randomization.

The following example shows this is not true.

Example 6. Again, take $\mathcal{C} = \{a, b\}$, $m = 4$, a preference relation with positive spillovers and $\mathcal{I}_1 = \{\{a\}, \{b\}\}$. Suppose that $(a, 2) \succ (b, 2)$, $(b, 3) \succ (a, 1)$ and that the most preferred case is $(b, 4)$, which happens in state $\tilde{x}$. Call $\hat{x}$ the state where every agent chooses a. Assume whatever imitation rule.

Under global information, the recurrent states are $\tilde{x}$ and $\hat{x}$. Note that 3 perturbations are required to move the system from $\hat{x}$ to $\tilde{x}$, while only 2 perturbations are required to move the system from $\tilde{x}$ to $\hat{x}$. By a mutation counting argument the unique stochastically stable state under global information is $\hat{x}$.

Since state $\tilde{x}$ is homogeneous within each information set under information structure $\mathcal{I}_1$, then by lemma 2 it is stochastically stable under $\mathcal{I}_1$. Therefore μ_g^* is not welfare superior to μ_1^*.

Example 6 is just a particular case of the following proposition, which is the unique analytical result I have been able to provide for a generic preference relation. Note in fact that there is no requirement on preferences, hence proposition 3 applies not only to cases of positive spillovers and negative spillovers but to any conceivable case.

Proposition 3. *If there is some agent who reaches the maximum of her preferences in no stochastically stable state under global information, then $\mathcal{I}_g$ is not welfare superior to any information structure $\mathcal{I}_1 \neq \mathcal{I}_g$.*

Proof. Suppose agent r does not reach the maximum of her preferences in any stochastically stable state under global information. Consider the information structure $\mathcal{I}_1 \neq \mathcal{I}_g$. Suppose one of her most preferred outcomes is (a, k), with $a \in I_i$. Take a state $\tilde{x}$ where agent r and other $k - 1$ agents choose a, and all other agents are grouped in an alternative out of the information set I_i (which is possible, since $\mathcal{I}_1 \neq \mathcal{I}_g$). State $\tilde{x}$ is homogeneous within each information set in $\mathcal{I}_1$, therefore by lemma 2 $\mu_1^*(\tilde{x}) > 0$, so proving that μ_g^* does not first-order stochastically dominate μ_1^* for agent r. $\square$

6 Conclusions

This paper has attempted to analyze the welfare implications of information in an evolutionary framework. When agents adopt behavioral rules based on imitation, is the disclosure of information about others' performances welfare enhancing?

The answer to such a question is "maybe" over finite time horizons, while in the long run with rare perturbations it tends to be "yes" in a pure coordination game (i.e. in the presence of pure positive spillovers) and "no" in a pure anti-coordination game (i.e. in the presence of pure negative spillovers).

This can be explained as follows. More information makes more agents align their choices in expectation. This favors coordination. More information can also provide more alternatives to copy, so dispersing the probability of the choice at the next time and spacing out agents' actual choices. This hinders coordination. However, this last effect disappears when considering the stochastically stable distributions, since the only recurrent states are homogeneous within each information set and hence no mixing effects are possible.

When a selection issue is introduced, by letting the preference relation to depend not only upon how many agents choose an alternative but also upon the kind of alternative itself, the answer to the former question is "maybe" even when working with the stochastically stable distributions. More information allows evolutionary forces to operate more freely. At a first glance this seems to favor global information over local information. However, evolutionary selection does not always prove superior to random selection. This fact can be thought of as an instance of the ambiguous relation between welfare and evolutionary efficiency.

As a project for future research, I consider potentially interesting further analyses about the welfare implications of information under behavioral rules different from imitation and in volatile setups of the model (multi-armed bandits).

References

G. M. Angeletos and A. Pavan. "Transparency of Information and Coordination in Economies with Investment Complementarities". *American Economic Review Papers and Proceedings*, 94:91–98, 2004.

G. M. Angeletos and A. Pavan. "Efficient Use of Information and Social Value of Information". *Econometrica*, 75:1103–1142, 2007.

R. J. Aumann. "Agreeing to Disagree". *The Annals of Statistics*, 4:1236–1239, 1976.

A. Ben-Shoham, R. Serrano, and O. Volij. "The Evolution of Exchange". *Journal of Economic Theory*, 114:310–328, 2004.

R. Cooper and A. John. "Coordinating Coordination Failures in Keynesian Models". *Quarterly Journal of Economics*, 103:441–463, 1988.

G. Ellison. "Basins of Attraction, Long Run Stochastic Stability, and the Speed of Step-by-Step Evolution". *Review of Economic Studies*, 67:17–45, 2000.

G. Ellison. "Learning, Local Interaction, and Coordination". *Econometrica*, 61:1047–1071, 1993.

G. Ellison and D. Fudenberg. "Rules of Thumb for Social Learning". *Journal of Political Economy*, 101:612–643, 1993.

D. Foster and H. P. Young. "Stochastic Evolutionary Game Dynamics". *Theoretical Population Biology*, 38:219–232, 1990.

M. I. Freidlin and A. D. Wentzell. *Random Perturbations of Dynamical Systems*. Springer, New York, 1984.

J. Hirshleifer. "The Private and Social Value of Information and the Reward to Inventive Activity". *American Economic Review*, 61:561–574, 1971.

M. Kandori, G. J. Mailath, and R. Rob. "Learning, Mutation and Long Run Equilibria in Games". *Econometrica*, 61:29–56, 1993.

S. Morris and H. S. Shin. "The Social Value of Public Information". *American Economic Review*, 92:1521–1534, 2002.

K. H. Schlag. "Why Imitate, and if so, How? A Boundedly Rational Approach to Multi-armed Bandits". *Journal of Economic Theory*, 78:130–156, 1998.

K. H. Schlag. "Which One Should I Imitate?". *Journal of Mathematical Economics*, 31:493–522, 1999.

F. Vega-Redondo. "The Evolution of Walrasian Behavior". *Econometrica*, 65:375–384, 1997.

H. P. Young. *Individual Strategy and Social Structure*. Princeton University Press, Princeton, 1998.

H. P. Young. "The Evolution of Conventions". *Econometrica*, 61:57–84, 1993.