Coordinating Regime Switches
Author(s): Christophe Chamley
Reviewed work(s):
Source: *The Quarterly Journal of Economics*, Vol. 114, No. 3 (Aug., 1999), pp. 869-905
Published by: Oxford University Press
Stable URL: http://www.jstor.org/stable/2586886
Accessed: 23/03/2012 14:46

# COORDINATING REGIME SWITCHES*

CHRISTOPHE CHAMLEY

The canonical model of strategic complementarities between individual actions, which exhibits multiple equilibria under perfect information, is extended with heterogeneous agents and imperfect information. Agents observe their own cost of action and the history of the levels of aggregate activity. The distribution of individual characteristics evolves through a random process, and individuals are rational Bayesians. Under plausible conditions, there is a unique equilibrium with phases of high and low activity and random switches. Applications may be found in macroeconomics and revolutions.

## INTRODUCTION

Payoff externalities between individual actions arise in structural models of search [Diamond 1982], and of investment with imperfect competition [Kiyotaki 1988; Murphy, Shleifer, and Vishny 1989]. These externalities generate strategic complementarities between individual actions and the possibility of multiple equilibria [Cooper and John 1988]. Beyond economics, strategic complementarities appear in models of social changes and revolutions [Kuran 1987, 1995].

Multiple equilibria suggest an analogy with the peaks and troughs of the business cycle. But the existence of multiple equilibria in itself does not provide a basis for fluctuations and cycles: why does one equilibrium arise rather than another? How do shifts between equilibria occur? This problem is investigated here in a canonical model with rational learning from the observation of aggregate activity. The model generates a unique equilibrium with random switches between regimes of high and low activity, in conformity with the empirical findings of Hamilton [1989], Filardo [1994], and Diebold, Rudebusch, and Sichel [1993], among others.

The macroeconomic context imposes two assumptions that have so far received little attention in the literature on coordination: payoff uncertainty and learning from history in a dynamic

* I am grateful to Olivier Blanchard and two referees for their remarkable help. Ricardo Caballero, Russell Cooper, Jonathan Eaton, Raquel Fernandez, Guy Laroque, Glenn Loury, and seminar participants at the NBER Summer Institute, Boston University, Princeton University, DELTA, the Massachusetts Institute of Technology, New York University, and CREST made useful suggestions. This is a thorough revision of "Coordination of Heterogeneous Agents in a Unique Equilibrium with Random Regime Switches," DELTA Discussion Paper 96-17, October 1996.



*The Quarterly Journal of Economics,* August 1999

model. I introduce these assumptions in an analytical framework that is based on the model of Cooper and John [1988]. The canonical form, while omitting important features of macroeconomic cycles, exhibits general properties that apply in other contexts of social changes. Agents make a zero-one decision in each period. Indivisible decisions play an important part in business cycles through the fluctuations of aggregate levels of lumpy expenditures. In politics, agents choose in a discrete set of candidates or regimes.

The imperfect information of agents is deeply related to their heterogeneity. Different agents have different costs of action (investment). All individuals make a decision simultaneously at the beginning of the period, and for each agent the payoff of action increases with the level of aggregate activity in that period. Individuals with relatively high cost have a positive net payoff only if the level of aggregate activity is sufficiently high.

The structure of the economy is defined by the distribution of individual costs (or parameters), and it evolves between consecutive periods by a random process with no discontinuity. A key assumption in the paper is that agents observe only the aggregate of individual choices (aggregate economic activity, percentage of votes for political parties) and their own private cost. As the structure of the economy evolves gradually, each agent uses as a Bayesian the information provided by the history of aggregate activities and his private cost to update his probability on the structure of the economy. His payoff of acting depends on the expectation on the level of aggregate activity during the period.

It will be shown that in an equilibrium, either most agents act, or they do not act. Therefore, the aggregate activity provides an observation on the tail of the distribution of agents (left or right), which in general conveys poor information on the entire distribution of costs. But the entire distribution matters for the possibility of switching from one level of aggregate activity to another. In this setting there will be a unique Bayesian equilibrium with regimes of low and high activity that alternate randomly.

A switch from, say, a low to a high regime occurs when the density of marginal agents increases (during the random evolution of the structure). Such an increase generates a higher level of activity, and by the multiplier effect of the positive externality, the switch to a high equilibrium. The higher activity also generates significantly more information. However, a switch will occur only for distributions of structural parameters such that under perfect information there would be a unique equilibrium. The sudden increase of information that is released by a switch will thus be compatible with a unique equilibrium.

A traditional method in models with multiple equilibria is to assume that agents coordinate on the one which is closest to the equilibrium in the previous period while the structure of the economy evolves smoothly [Goodwin 1951; Cooper 1994]. When an equilibrium of this "type" disappears, the economy jumps to an equilibrium of a different type. This rule of thumb is ad hoc and not satisfactory,[1] but it generates an aggregate behavior with hysteresis that is similar to the equilibrium in the present model. However, the two approaches lead to different conclusions for policy.

The paper is organized as follows. A variation of the canonical model of Cooper and John [1988] is introduced in Section I. Section II briefly reviews the case of perfect information in which there may be two (stable) equilibria with low and high activity, respectively. Section III is the core of the paper and analyzes the mechanisms by which there is a unique equilibrium with random switches between regimes of high and low activity. This property depends on two types of restrictions on the parameters of the model: first, in a state of low (high) activity, there is a strictly positive probability that no coordination is feasible at a high (low) level of activity; second, the degree of agents' heterogeneity cannot be too small. I first present informally the workings of the unique equilibrium. The existence and uniqueness are then provided analytically under sufficient conditions that are fairly strong, for simplicity. The main technical arguments are explained in the text, but the proofs are in the Appendix.

The properties of the model are then shown to be robust under a partial relaxation of the main assumptions using some intuitive arguments and numerical simulations in Section IV. Assumptions and results are also compared with those of Carlsson and Van Damme [1993] and others.

Section V is devoted to an application to political changes with a dynamic version of Kuran's [1995] model, and to the final comments.

1. It might be justified by inertia, a concept that remains to be investigated in the context of strategic uncertainty and delays [Morris 1995]. In the present model, however, inertia is ruled out because agents have only one opportunity to make a decision and cannot delay.

## I. The Model

There is a continuum of agents, and time is discrete. The population is new in each period, and individuals live one period. Following the discussion in the Introduction, each agent has to make a one or zero decision, whether to *act* or *not act.* An agent who does not act has a zero payoff. The payoff of acting, $u$, is the difference between a *payoff externality* $v$, and a *private cost* $c$: $u = v - c$. The term $v$ is the same for all agents while the private cost $c$, which can be negative (when private benefits exceed the private cost), is specific to an agent and defines him. The model is built on (i) the specification of the payoff function $v$; (ii) the distribution of agents with its law of random evolution; (iii) the information-generating process.

The payoff externality $v$ is a positive increasing function of the mass $Y$ of agents who act in the same period. The externality generates a strategic complementarity between individual actions. In order to simplify the analysis and without loss of generality, it will be assumed that the function $v$ is linear. With a proper normalization, $v(Y) = Y$. Under uncertainty, agents maximize the expected value of their payoff.

Agents are characterized by their individual cost $c$. The distribution of individual costs is assumed to be rectangular as represented in Figure Ia, and is characterized by the density function $f$ with

$$(1) \qquad f_x(c) = \begin{cases} \beta & \text{for } -b \leq c \leq x \text{ and } x + \sigma \leq c \leq B, \\ \alpha + \beta & \text{for } x < c < x + \sigma, \end{cases}$$

where $\sigma$ is constant, $x$ is a random variable, and $-b$, $B$ are the boundaries of the distribution, ($b > 0, B > 0$). All agents know the values $\alpha$, $\beta$, and $\sigma$, but the variable $x$ is not directly observable. The distribution has the important property that the value of $x$ is unknowable from an observation of the cumulative distribution function $F_x(c)$ when $c$ is in one of the tails of the distribution.

For the existence of multiple equilibria, $\alpha$ and $\beta$ are such that

$$0 < 1 - \beta < \alpha.$$

The population can be viewed as the sum of two subpopulations. The first has a uniform density of costs equal to $\beta$ on the interval $[-b, B]$. The second group has a uniform density $\alpha$ on the interval $I_x = (x, x + \sigma)$ that is contained in the interval $[-b, B]$. This group will be called the "cluster." Its definition is only an analytical

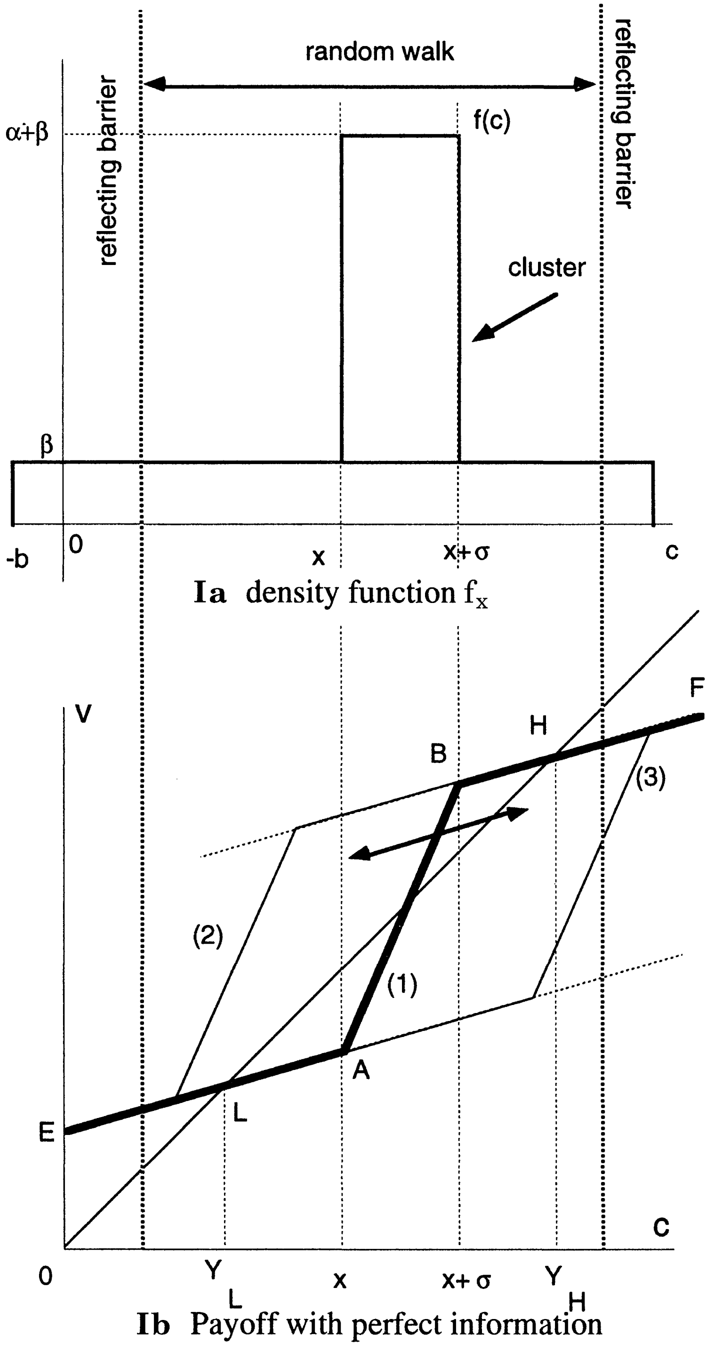


FIGURE I
Structure and Equilibria under Perfect Information

device: agents with $c \in I_x$ can be individually assigned to the cluster or to the uniform distribution.

The function $f$ in (1) is chosen as a stylized representation of an economy with some clustering and some heterogeneity: with no clustering and a diffuse distribution, there would be no critical mass of similar agents who can act together and generate multiple equilibria; but some heterogeneity will be essential for the argument of the paper. A generalization of the density $f$ toward "smoother" function will be discussed in subsection IV.E.

The realization of $x$ in period $t$, $x_t$, defines the structure of the economy in period $t$. A critical assumption here is that the structure of the economy does not jump between periods, but evolves only gradually. (Following Leibnitz, "Natura non facit saltus.") Accordingly, $x$ evolves from one period to the next in a one-step discrete random walk. In order to keep the model bounded, the random walk is subject to two reflecting barriers:[2] there are two values $\gamma$ and $\Gamma$ such that for all $t$, $\gamma \leq x_t \leq \Gamma$, $(-b < \gamma < \Gamma < B - \sigma)$.

The values of $x_t$ are in the discrete grid $\omega = \{\omega_k\}$, with $\omega_k = \gamma + (k-1)\epsilon$, $k = 1, \ldots, K$, $\gamma > 0$, $\epsilon > 0$. For convenience, the ratio between the width of the cluster $\sigma$, and width of the grid $\epsilon$ is an even integer. The value of $\epsilon$ will be small in a sense specified later. It will not have to be infinitesimal, however.

Let $p$ be a positive parameter, $0 < p \leq 1/3$. Denoting a probability by $P$, the random evolution of $x_t$ is defined by the following equations:

$$(2)\quad \begin{cases} \text{if } \omega_1 < x_t < \omega_K, & P(x_{t+1} = x_t + 1) = P(x_{t+1} = x_t - 1) = p, \\ & P(x_{t+1} = x_t) = 1 - 2p; \\ \text{if } x_t = \omega_1, & P(x_{t+1} = x_t + 1) = p, \quad P(x_{t+1} = x_t) = 1-p; \\ \text{if } x_t = \omega_K, & P(x_{t+1} = x_t - 1) = p; \quad P(x_{t+1} = x_t) = 1 - p. \end{cases}$$

The assumption that $p \leq 1/3$ is reasonable: the random evolution of $x_t$ in the previous equations can be viewed as the discrete specification of a smooth diffusion process in continuous time. In such a process the distribution of $x_t$, which evolves from an initial value $x_0$, is always hump-shaped. For the discrete formulation, this property is satisfied only when $p \leq 1/3$. The asymptotic distribution of $x_t$ does not depend on the value of $p$, however, and

2. The assumption of reflecting barriers also embodies a regression of $x_t$ toward its mean.

numerical simulations show that the properties of the model may hold when $p > 1/3$.

## II. Perfect Information: Multiple Equilibria

In this section it is assumed that agents know perfectly the value of $x$ and therefore the distribution of costs in the economy (as in other studies on coordination [Bryant 1987; Cooper and John 1989; Cooper 1994]). If the payoff of an agent with cost $\hat{c}$ is positive, it is also positive for any agent with a cost $c < \hat{c}$. Therefore, in an equilibrium the *acting set,* which is the set of costs of all acting agents, is an interval of the form $[-b, c^*]$, and the mass of acting agents is the value of the cumulative distribution function $F_x(c^*)$, which depends on the realization of the parameter $x$. An equilibrium is characterized by a value $c^*$ such that $c^* = F_x(c^*)$, and $c^*$ is the highest cost of acting agents.

Given the particular realization of $x$ in Figure I, there are two stable equilibria with levels of aggregate activity $Y_L$ and $Y_H$:

$$Y_L = b\beta/(1 - \beta)\,, \qquad \text{and} \qquad Y_H = (b\beta + \alpha\sigma)/(1 - \beta)\,.$$

The two stable equilibria are represented by the points $L$ and $H$ in Figure Ib. For the existence of these equilibria, it will be assumed throughout the paper that

$$(3) \qquad b > 0, \qquad \text{and} \qquad B > (\beta b + \alpha\sigma)/(1 - \beta)\,.$$

These inequalities will ensure that there is a positive mass of agents who always act and another who never act. By assumption, the variations of the cluster are such that for some realizations of $x$ all agents in the cluster act (in an equilibrium), while for some others none of these agents act. The necessary and sufficient condition for the existence of these regimes is

$$(4) \qquad \omega_1 < Y_L, \qquad \text{and} \qquad \omega_K + \sigma > Y_H.$$

These inequalities will play an essential role in the paper. We will show that a switch between high and low levels of activity will occur when $x$ takes a value on the grid points near $Y_L$ and $Y_H$. We therefore single out this point by the following notation. The grid points $\omega_M$ and $\omega_N$ are nearest $Y_L$ and $Y_H$ in the sense that

$$(5) \qquad \omega_M < Y_L < \omega_{M+1}, \qquad \text{and} \qquad \omega_{N-1} < Y_H - \sigma < \omega_N.$$

Depending on the realization of $x_t$, there may be one or two stable equilibria with levels of activity $Y_L$ and $Y_H$, respectively. For the sake of clarity, these equilibria are characterized in the next result which needs no proof.

PROPOSITION 1. Under perfect information about $x_t$, there are two possible levels of activity in a stable equilibrium, $Y_L$ and $Y_H (Y_L < Y_H)$, respectively.
The structure of equilibria in any period $t$ is characterized as follows.

- If $\omega_1 \leq x_t < Y_L$ (which is equivalent to $\omega_1 \leq x_t \leq \omega_M$), the average cost of action is relatively low, and there is one equilibrium: it has a high level of activity $Y_H$, and all agents in the cluster act.
- If $Y_L \leq x_t \leq Y_H - \omega$ ($\omega_{M+1} \leq x_t \leq \omega_{N-1}$), the average cost is in an intermediate range. There are two equilibria with levels of activity $Y_L$ and $Y_H$, respectively. In the high equilibrium all agents in the cluster act. In the low equilibrium no agent in the cluster acts.
- If $Y_H - \sigma < x_t \leq K$ ($\omega_N \leq x_t \leq \omega_K$), the average cost is relatively high, and there is one equilibrium: it has a low level of activity $Y_L$, and no agent in the cluster acts.

The three cases are represented in Figure Ib by the payoff functions (2), (1), and (3), respectively. The term "stable" in the proposition is standard and does not need to be justified formally. Given the random walk property of $x_t$, the probability that any of the possible situations will arise eventually is equal to one. Figure Ib illustrates the relation between the present model and the canonical model of Cooper and John [1989].

## III. IMPERFECT INFORMATION: THE EQUILIBRIUM

We now assume that agents have imperfect information on $x$ (and therefore on the distribution of costs) at the beginning of the initial period (period 0). In that period nature chooses an initial value for that parameter, $x_0$, according to a probability distribution $\pi_0 = (\pi_{1,0}, \ldots, \pi_{K,0})$ on the grid $\omega = \{\omega_1, \ldots, \omega_K\}$. This initial distribution, which will satisfy some assumption later, is known to all agents. The value of $x_t$ is never observable directly, but the equations of the random evolution (2) are known to all agents. For any period $t \geq 2$, the history of aggregate activities for the past periods, $h_t = \{\pi_0, Y_1, \ldots, Y_{t-1}\}$, is common knowledge.

Each agent has some additional private information through his own cost $c$, and combines this information with the common knowledge to form his own probability assessment about the distribution of $x_t$. Private costs are not publicly observable. Each agent acts if and only if his expected payoff is positive.

### III.A. Equilibrium and the Value Function

The set of costs of the acting agents in period $t$ is called the *acting set* in period $t$, and is denoted by $A_t$. Given a realization $x_t$ which defines a particular cost distribution, aggregate activity in period $t$ is equal to the mass of agents in $A_t$:

$$Y_t = \mu_{x_t}(A_t),$$

where $\mu_{x_t}(A_t)$ is the Lebesgue measure of $A_t$ for the population distribution associated with the realization $x_t$. The net payoff of action for an agent with cost $c$ is the difference between his expected value of the mass of agents in the acting set and his cost:

$$u(A_t, h_t, c) = E[Y_t|A_t; h_t, c] - c = E[\mu_{x_t}(A_t)|h_t, c] - c,$$

where the expectation is taken on $x_t$, conditional on the information $(h_t, c)$. An *equilibrium acting set* $A_t$ in period $t$ is defined such that

$$c \in A_t \qquad \text{if and only if} \qquad u(A_t, h_t, c) \geq 0.$$

Any agent can compute the payoff and the strategy of any agent with information $(h_t, c)$. Hence all agents can agree on the same acting set. However, note that agents have different private estimates on $x_t$ (through the observation of their own cost), and therefore different estimates of the mass of agents in the acting set.

At the end of period $t$ the observation $Y_t$ is used to update in Bayesian fashion the commonly known distribution $\pi_t$ (of $x_t$ *for period* $t$) to $\tilde{\pi}_t$. The probability assessment $\pi_{t+1}$ of $x_{t+1}$ is then obtained from $\tilde{\pi}_t$ by application of the transition equations (2). We will say that in this case the probability assessment $\pi_{t+1}$ is consistent with $A_t$ and the history $h_{t+1}$ of the economy. The process is repeated for any period. We have therefore the following definition of a Bayesian equilibrium.

DEFINITION 1. An equilibrium is defined by a sequence of acting sets $\{A_t\}_{t \geq 1}$ and probability assessments $\pi_t$ such that for any realization of $\{x_t\}$,

$$c \in A_t \qquad \text{if and only if} \qquad c \leq E[Y_t|A_t; h_t, c] = E[\mu_{x_t}(A_t)|h_t, c],$$

and for any period $t$, $\pi_t$ is consistent with the history $h_t$ and the acting sets in previous periods.

If beliefs $\pi_t$ are consistent with the history $h_t$ for any period, the information of $h_t$ is equivalent to the probability distribution $\pi_t$. The information $\{h_t,c\}$ is summarized by $\{\pi_t,c\}$. We have seen that under perfect information, an equilibrium acting set is an interval $[-b, c^*]$ with $c^*$ defined by the equation $F_x(c^*) = c^*$. This equation is now generalized to the case of imperfect information where agents have different expectations depending on their own cost. A useful tool of analysis will be the *value function* at the point $c$ which is defined as the expectation for an agent with cost $c$ of the mass of agents with cost less than $c$.

DEFINITION 2. The value function in period $t$ is defined as $V_t(c) = E[F_x(c)|h_t,c]$, where $F_x$ is the cumulative function associated with $x$.

The value function $V$ is now used to construct two sets of agents who find it iteratively dominant to act or to not act, respectively. Under a simple condition on $V$ these two sets will cover the entire range of costs, and the existence and uniqueness of the Nash equilibrium will follow immediately.

Consider first an agent with negative cost. If he acts, his payoff is positive no matter what the others choose to do. Acting is a dominant strategy and strictly dominant if $c < 0$. The mass of acting agents is therefore bounded below by $\beta b > 0$. If the agents with negative cost choose their dominant strategy, then to act becomes the dominant strategy for the agents with a cost that is positive and less than $\beta b$. We can iterate this step by a well-known method. The argument is formalized by the construction of a sequence of sets $A^k$, such that

$$(6)\qquad \begin{cases} A^0 = [-b,0], \\ A^k = \{c | E_x[\mu_x(A^{k-1}) > c | \pi,c]\}, k \geq 1, \\ \overline{A} = \cup_{k=0}^{k=\infty} A^k. \end{cases}$$

Expectations depend on the history of the economy, but the period subscript is omitted for simplicity. For an agent with cost $c \in \overline{A}$, acting is, by definition, *iteratively dominant.* Note that the sequence of set $A^k$ is increasing. In the same way, for the agents with a cost in the interval $(\beta(b + B) + \alpha\sigma, B]$ nonacting is a dominant strategy. Denoting the complement of a subset $D$ with respect to $[-b, B]$ by $\tilde{D}$, one defines the sequence of sets as

$$(7) \qquad \begin{cases} D^0 = (\beta(b + B) + \alpha\sigma, B], \\ D^k = \{c \,|\, E_x[\mu_x(\tilde{D}^{k-1}) < c \,|\, \pi, c]\}, \ k \geq 1, \\ \overline{D} \ = \cup_{k=0}^{k=\infty} D^k. \end{cases}$$

For an agent with cost $c \in \overline{D}$, not acting is iteratively dominant. The sequence of set $D^k$ is increasing. The sets of agents with an iteratively dominant strategy are related to the value function $V$ by the following result.

PROPOSITION 2.

a. Suppose that there is some $c^*$ such that $V_t(c) > c$ for $c \in [-b, c^*)$. Then for any agent with cost $c < c^*$, acting is iteratively dominant in period $t$.
b. Suppose that there is some $c^*$ such that $V_t(c) < c$ for $c \in (c^*, B]$. Then for any agent with cost $c > c^*$, not acting is iteratively dominant in period $t$.

The proof of the result (in the Appendix) follows the intuition that is provided by the construction of the sets of dominant strategies $A^k$ and $D^k$. From this Proposition we immediately have a sufficient condition for the existence of a unique equilibrium acting set in any period.

PROPOSITION 3. Assume that the function $V_t$ satisfies the following condition: there exists $c_t^*$ with $V_t(c) > c$ for $c < c_t^*$, and $V_t(c) < c$ for $c > c_t^*$. Then, there exists a unique equilibrium acting set in period $t$: $A_t = [-b, c_t^*]$.

This result provides the criterion for the demonstration of the uniqueness of any equilibrium in subsection III.C.

*III.B. An Informal Presentation of the Equilibrium*

Suppose that in period 0, $x_0 = \omega_N$, and that $x_0$ is known with perfect information: this is the lowest value on the grid such that under perfect information, there would be a unique equilibrium with low activity, $Y_L$. The value function in period 0 is identical to the cumulative distribution function (c.d.f.) $F_{x_0}$. Its graph is represented by $F_0$ in Figure II (line *LDEC*), and has the same shape as a c.d.f. of type 3 in Figure I. The cluster in period 0 is represented by the thick line on the horizontal axis. We will see later that the assumption of perfect information for period 0 is relevant in the setting of imperfect information. (Agents will learn

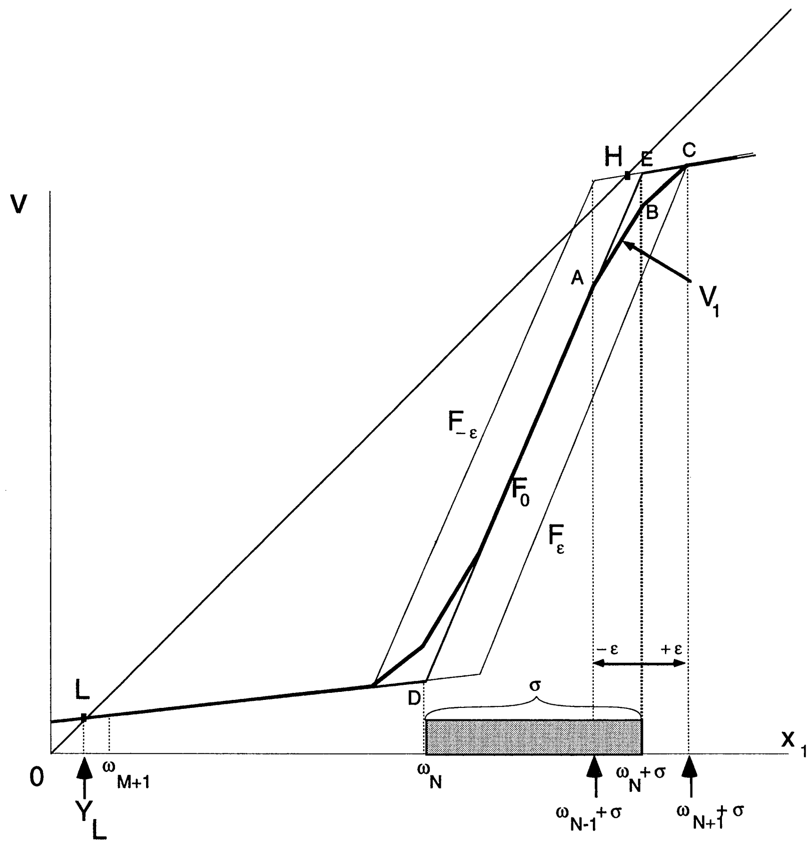


FIGURE II

$x$ in any period when a switch of regime occurs). We proceed now for any period $t \geq 1$ with the assumption that agents observe only the level of aggregate activity.

A key difference between the cases of perfect and impefect information arises already in period 1. Between periods 0 and 1, the cluster has either stayed, moved to the right by $\epsilon$, or moved to the left by $\epsilon$. Suppose first that the cluster has moved to the left. Under perfect information $x$ is known, and the value function is the c.d.f $F_{-\epsilon}$ which is represented in Figure II: there are two equilibria $H$ and $L$, respectively; the graph is essentially the same as in case (1) of Figure Ib. Under imperfect information, the move of $x$ is not observable, and the value function $V_1(c)$ is an average of the three c.d.f. $F_{-\epsilon}$, $F_0$, and $F_\epsilon$. The probability weights depend on the cost of $c$ of the agent, which is private information. The key insight is that because the value functions under perfect information $F_{-\epsilon}$, $F_0$, and $F_\epsilon$, are concave near $H$, the average is smaller than the middle function $F_0$, for which there is no equilibrium with high activity. Hence, there is no equilibrium with high activity under impefect information when $x$ has moved to the left in period 1.[3]

The previous argument applies only if the probability of a leftward shift is not too high for the subjective estimate of an agent with cost $c$. For such an agent the subjective probabilities that the cluster has moved to the left, stayed, or moved to the right, respectively, are obtained in two steps: first, the transition rule of equation (2) generates a common knowledge probability distribution on the three possible cases $\{F_{-\epsilon}, F_0, F_\epsilon\}$. That distribution, $(p, 1-2p, p)$ is symmetric around the middle position $F_0$. Second, the information of one's cost $c$ is used to update this common distribution. One needs to consider only the values of $c$ for which the c.d.f. can change. Near the point $H$, these values are in the interval $I = [\omega_{N-1} + \sigma, \omega_{N+1} + \sigma]$. An agent with cost $c$ tends to increase his estimate that he is in the cluster; that is, that $x + \sigma > c$. Therefore, the individual update puts more weight on rightward than leftward moves. Since $F_{-\epsilon} < F_\epsilon$, for each agent with $c \in I$, the expected value $E[F(c)]$, is smaller than the average under the symmetric common knowledge distribution. That average is itself lower than $F_0$ because of the concavity of the c.d.f. The argument is illustrated in Figure II by the line $ABC$ which is a "smoothing" of $V_0$.[4] The only equilibrium level of activity is the low one, at $Y_L$.

The argument is now extended for $t > 1$. Suppose that for all previous periods $\tau < t$: (i) the maximum cost of acting agents $c_\tau^*$ is smaller than $\omega_{M+1}$, which is the grid point immediately to the right of $Y_L$; (ii) the realization $x_\tau$ is strictly greater than $c$. These properties will later define the *low regime*. In this regime the level

3. There is obviously no difference between the outcomes under perfect and imperfect information if $x$ shifts to the right.

4. As shown in Figure II, the value function $V_1(c)$ is also different from $V_0(c)$ for $c \in (\omega_{N-1}, \omega_{N+1})$, but this effect does not change the equilibrium set.

of aggregate activity is equal to $b + \beta c_{\tau}^{*}$, and reveals only that $x_{\tau} \geq c_{\tau}^{*}$. This observation is compatible with a wide range of values $x_{\tau} \geq \omega_{M+1}$.

As long as the economy stays in a low regime, the common knowledge distribution of $x_{\tau}$ which is inferred from the history of aggregate activities spreads out gradually to the entire interval $[\omega_{M+1}, B]$. The evolution of $\pi_{\tau}$ is represented in Figure III. Note how the distribution puts stronger weights on relatively high costs: when agents observe no action from the cluster, they revise their probabilistic position of the cluster to the right, toward higher cost values. The asymptotic distribution $\pi_{\infty} = \text{Lim}_{t \to \infty} \pi_t$ has a simple sine form which is derived in the Appendix.

Each agent with cost $c$ combines $\pi_t$ with the information of his private cost to have his subjective distribution $\tilde{\pi}_t(c)$, and the value function at the point $c$ is the average of the values of cumulative

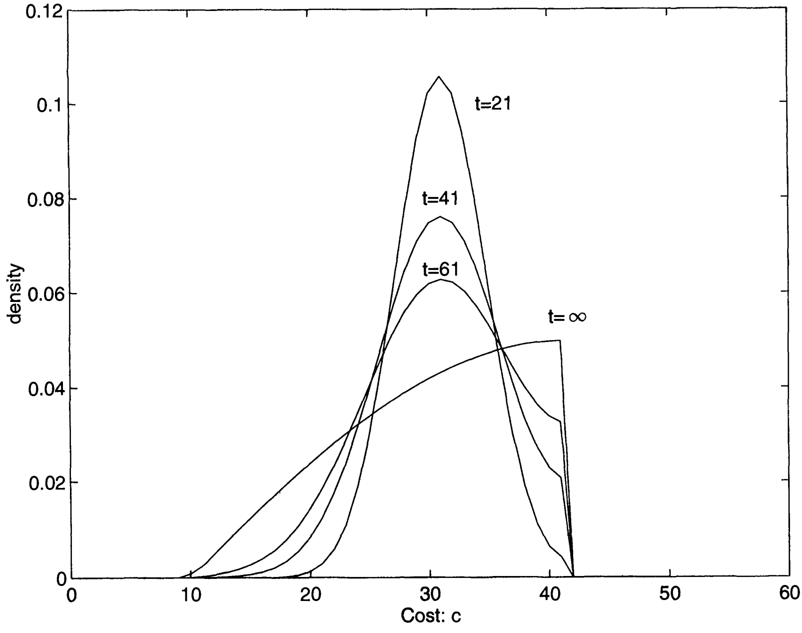


FIGURE III
Evolution of Common Knowledge in the Low Regime
Parameters: $\epsilon = 1$, $\omega_M = 10 = Y_L - 0.5$, $\omega_N = 31 = Y_H - \sigma + 0.5$, $\omega_K = 41$, $\sigma = 10$, $\beta = 0.3$, $\alpha = 2.1$, $p = 1/3$.

distribution functions:

$$V_t(c) = \sum_k \tilde{\pi}_{k,t}(c) F_{\omega_k}(c),$$

where $\tilde{\pi}_{k,t}(c)$ is the probability for agent $c$ that $x_t = \omega_k$. Two cases should be distinguished here. In the first, which will be the main one in the paper, the level of heterogeneity is not too small: one's individual cost is not too informative on the costs of others and the updated distribution $\tilde{\pi}_t(c)$ is not too different from $\pi_t$. The skewness of $\pi_t$ to the right (as shown in Figure III) holds for each subjective distribution $\tilde{\pi}_t(c)$ and depresses the value information $V_t(c)$. Figure IV represents the evolution through time of the value function $V_t$ in a low regime and conditional on realizations $x_\tau \geq \omega_{M+1}$ for all $\tau \leq t$. The contrast with the value functions in Figure I and its multiple equilibria illustrates the impact of imperfect information. One can also verify visually that the sufficient condition in Proposition 3 applies and that there is a unique equilibrium.

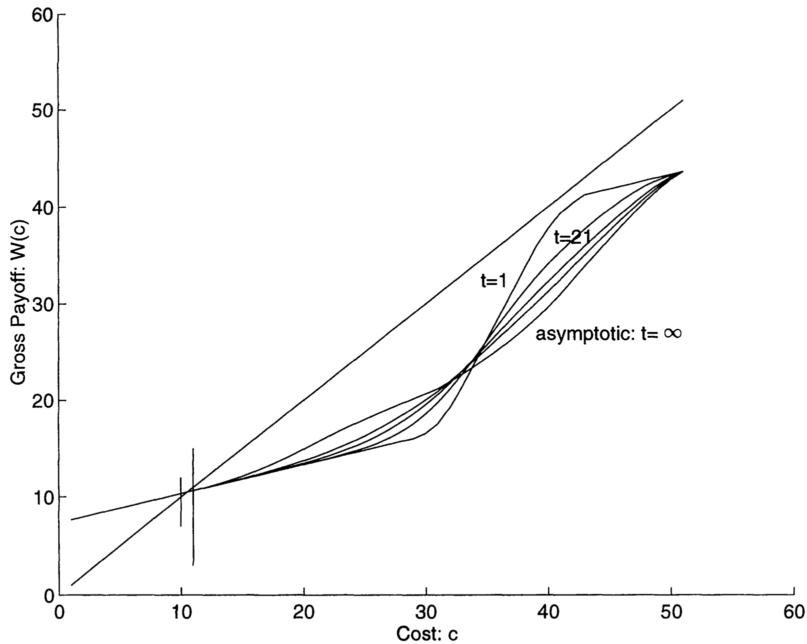


FIGURE IV
The Evolution of the Value Function (Same Parameters as in Figure III)

For the second case where heterogeneity is small, one's cost $c$ conveys significant information on the cost of others, and the skewness property may not be maintained in the updating from $\tilde{\pi}_t$ to $\tilde{\pi}_t(c)$. In the extreme case of vanishing heterogeneity, any agent with a cost $c$ not too close to $Y_L$ (and this is an important condition), believes that half the other agents have a cost lower than $c$. If he expects them to act, he will also act even if his cost is fairly large. This case will be discussed further in subsection IV.B.

Things change dramatically in the first period for which the realization of $x_t$ is smaller than the maximum cost of acting agents $c_t^*$. At the *end* of that period, agents observe that some agents of the cluster have acted. Because $x_t$ moves by one step between periods, the exact value of $x_t$ can be identified. At the beginning of period $t + 1$, agents have perfect information on the value of $x_t$ (in the previous period). This situation is isomorphic to the one we considered in period 1 in which agents knew the value of $x_0$. Similar arguments can be used, mutatis mutandis: there is now a unique equilibrium with high activity $Y_H$. A high regime begins with aggregate activity at or near the high level $Y_H$. It continues as long as the realizations of $x_t$ are such that $x_t \leq \omega_{N-1}$. At the end of the first period $t$ for which $x_t = \omega_N$, the situation is the same as in period 0: the economy plunges again in a phase of low activity. The turning point from high to low activity is identical to the one in period 1. This property of the turning points justifies the assumption of perfect information that was made for period 0.

### *III.C. Formal Results*

Recall first the definitions of the three ranges for $x$: if agents can observe $x$, there is a unique equilibrium with high activity when $x \in [\omega_1,\omega_M]$, multiple equilibria if $x \in (\omega_M,\omega_N)$, and a unique equilibrium with low activity if $x \in [\omega_N,\omega_K]$. The informal discussion has highlighted the importance of two assumptions. First, there must be some positive probability that $x$ is in the range of costs where there is only one equilibrium under perfect information. In this section it will be assumed that these two ranges are wider than the range with multiple equilibria $(\omega_M,\omega_N)$. This assumption is stronger than necessary as will be shown in Section IV, but it is introduced to keep the technicalities relatively simple.

ASSUMPTION 1. $\omega_M - \omega_1 \geq \omega_N - \omega_M$, and $\omega_K - \omega_N \geq \omega_N - \omega_M$.

The second assumption must define a lower bound on heterogeneity. This lower bound will be expressed here by the next assumption.

ASSUMPTION 2. $\alpha < 2(1 - \beta)$.

As for Assumption 1, a weaker condition would be sufficient for the properties of the paper, but Assumption 2 is introduced for analytical convenience. Its relaxation will also be examined in Section IV. The economy will always be in one of two regimes with "low" or "high" activity. In these regimes the cutoff cost $c_t^*$ will be near one of its values under perfect information with activity $Y_L$ or $Y_H$. The regimes are defined formally as follows.

DEFINITION 3.

In a low regime an agent acts if and only if his cost is smaller than $c_t^*$, where $Y_L \leq c_t^* < \omega_{M+1}$.

In a high regime an agent acts if and only if his cost is smaller than $c_t^*$, where $\omega_{N-1} < c_t^* \leq Y_L$.

In any of the two regimes, the cutoff value $c_t^*$ is within $\epsilon$ of $Y_L$ or $Y_H$ (using the definitions of $\omega_M$ and $\omega_N$ in (5), and $\omega_{i+1} - \omega_i = \epsilon$). Suppose that the economy is in a low regime in period $t$. There are two possible outcomes at the end of the period.

- If $x_t > c_t^*$, no agent in the cluster has acted in period $t$. The level of activity is equal to $Y_t = \beta(b + c_t^*)$, which is independent of $x_t$. The observation of $Y_t$ is compatible with any value $x_t > c_t^*$ and reveals that $x_t > c_t^*$.
- If $x_t < c_t^*$, some agents in the cluster act in period $t$ (who are in the interval[5] $(x_t, c_t^*)$. The observation of aggregate activity $Y_t = \beta(b + c_t^*) + \alpha(c_t^* - x_t) > c_t^*$ reveals perfectly the value of $x_t$.

As discussed in the informal presentation and illustrated in Figure III, the common knowledge distribution $\pi_t$ puts more weight on higher costs. Three specific features of this distribution are characterized in the definition of Property $\mathscr{H}$ (which stands for high costs):

PROPERTY $\mathscr{H}$.

(i) $\pi_{i,t} = 0$ for $i < M$.

(ii) $\pi_{N-i,t} \leq \pi_{N+i,t}$ for $N + i \leq K$.

(iii) $\pi_{i,t} \leq \pi_{i+1,t}$ for $M \leq i < N$.

5. If $x_t < Y_L - \sigma/(1 - \beta) < c_t^*$, there would be no perfect information, but this case cannot occur before the previous one since $x_t$ moves by at most $\epsilon$ between periods. Once $x_t < c_t^*$, the regime of low activity ends.

Part (i) rules out values of $x$ smaller than $\omega_M$; part (ii) defines a skewness such that for two grid points equidistant from $\omega_N$, the right one has higher probability; according to (iii), the density of $x$ is increasing in the interval from $\omega_M$ to $\omega_N$. the converse of Property $\mathscr{H}$ will be used in a high regime.

PROPERTY $\mathscr{L}$.

(i) $\pi_{i,t} = 0$ for $i > N$.
(ii) $\pi_{i,t} \geq \pi_{i+1,t}$ for $M \leq i < N$.
(iii) $\pi_{M-i,t} \geq \pi_{M+i,t}$ for $M - i \geq 1$.

The skewness in Property $\mathscr{H}$ or $\mathscr{L}$ is much weaker than what is observed in Figure III with numerical simulations. But it is technically easy to handle because it is self-reproducing in a low or a high regime.

LEMMA 1. In a low regime, if $\pi_t$ satisfies Property $\mathscr{H}$ in an equilibrium, and $x_t > Y_L$, then $\pi_{t+1}$ satisfies Property $\mathscr{H}$. In a high regime, if $\pi_t$ satisfies Property $\mathscr{L}$ in an equilibrium, and $x_t < Y_H$, then $\pi_{t+1}$ satisfies Property $\mathscr{L}$.

In order to simplify the analysis, we assume in the rest of the section that the values $Y_H$ and $Y_L$ are at the middle of their respective grid intervals:[6]

$$Y_L = (\omega_M + \omega_{M+1})/2, \text{ and } Y_H = (\omega_{N-1} + \omega_N)/2.$$

The main result of this section shows that if the grid is sufficiently fine ($\epsilon$ significantly small), there is a unique equilibrium.

PROPOSITION 4.

Under Assumptions 1 and 2 there exists $\epsilon^* > 0$ such that if $\epsilon < \epsilon^*$, and if $\pi_0$ satisfies the property $\mathscr{H}$ or $\mathscr{L}$, then the economy has a unique equilibrium. In any period $t$, if $\pi_t$ satisfies property $\mathscr{H}$, the economy is in a low regime: $Y_L \leq c_t^* < \omega_{M+1}$.

- If $x_t > Y_L$, Property $\mathscr{H}$ is satisfied in period $t + 1$.
- If $x_t < Y_L$, the value of $x_t$ is revealed perfectly at the end of period $t$, and the economy switches to a high regime in period $t + 1$ in which $\pi_{t+1}$ satisfies the property $\mathscr{L}$.

If $\pi_t$ satisfies Property $\mathscr{L}$, the economy is in a high regime, and the rules are similar, mutatis mutandis.

Note that the restriction of Property $\mathscr{H}$ for the initial period 0 is fairly mild since it is satisfied by a uniform distribution to the

6. This assumption was relaxed in the discussion paper version.

right of $Y_L - \epsilon$. The rest of this section is devoted to the proof for the low regime, using the steps of the arguments that were introduced in the informal presentation. We will use the function $W_t(c)$, which is defined as the expectation (for an individual $c$), of the mass of agents in the cluster "left" of $c$. Instead of Property 1, we will sometimes use the following result.

LEMMA 2. $V_t(c) \leq c$ if and only if $W_t(c) \leq (1 - \beta)(c - Y_L)$, with $W_t(c) = E[\min(\alpha\sigma, \alpha \max(c - x_t, 0) | h_t, c]$.

The higher the individual cost $c$, the higher the expectation $W_t(c)$ must be for a positive payoff of action. We have seen in subsection III.B that the property $\mathscr{H}$ puts a bias on $x$ toward higher values. This effect lowers $W_t(c)$. But each agent with cost $c$ updates the distribution from $\pi_t$ to $\pi_t(c)$ by Bayes' rule:

$$\frac{\pi_{i,t}(c)}{\pi_{j,t}(c)} = \left(\frac{\pi_{i,t}}{\pi_{j,t}}\right)\left(\frac{\beta + \alpha\gamma(\omega_i, c)}{\beta + \alpha\gamma(\omega_j, c)}\right), \tag{4}$$

$$\text{with} \quad \gamma(\omega_i, c) = \begin{cases} 1 & \text{if } c - \sigma \leq \omega_i \leq c, \\ 0 & \text{if } \omega_i < c - \sigma \text{ or } \omega_i > c - \sigma. \end{cases}$$

Because of the clustering, each agent is biased toward putting himself in the cluster and believing that $x \in [c - \sigma, c]$. The individual updating shifts $W_t(c)$ toward the value $\alpha\sigma/2$.

The first part of the proof is to show that for any cost $c > (Y_L + Y_H)/2$, $V_t(c) < c$. There are two cases, for agents with $c \geq \omega_N + \sigma/2$ and $c < \omega_N + \sigma/2$, respectively.

- Consider first an agent with $c > \omega_N + \sigma/2$ and a fictitious experiment such that the cost distribution in period 0, $\pi_0$, is the perfect information that $x_0 = \omega_N$. With this information an agent with cost $c$ would infer that his cost is in the upper half of the cluster. From Property $\mathscr{H}$ the distribution $\pi_t$ is the result of a diffusion from the initial distribution $\pi_0$, with a skew to the right. This property reinforces agents' belief that at least half the cluster has a higher cost than $\omega_N$.
- Consider now an agent with cost $c \leq \omega_N + \sigma/2$. With the distribution $\pi_0$ of perfect information $x_0 = \omega_N$, he sees that less than half the cluster has a cost lower than his own: $W_0(c) \leq \alpha\sigma/2$. With the bias of $\pi_t$ toward high costs, this inequality is maintained in the low regime after period 0:

$W_t(c) \leq \alpha\sigma/2$ and $V_t(c) \leq \beta(b + c) + \alpha\sigma/2$. If $c > (\beta b + \alpha\sigma/2)/(1 - \beta) = (Y_L + Y_H)/2$, $V_t(c) < c$.

These two arguments apply when the distribution of $x$ can spread sufficiently to the right of the critical value $\omega_N$, which is guaranteed by Assumption 1. They are used for the following result which is proved in the Appendix.

LEMMA 3. Under Assumption 1, if $\pi_t$ satisfies the property $\mathscr{H}$, $V_t(c) < c$ for any $c > (Y_L + Y_H)/2$.

The result holds because the possibility of high costs (case (3) in Figure I) depresses the value function. The strength of this effect is ensured by Assumption 1. For agents with a cost higher than the middle of the interval $[Y_L, Y_H]$, the value function is smaller than $c$. The highest cost of acting agents must therefore be smaller than $(Y_L + Y_H)/2$.

We now turn to agents with a cost smaller than $(Y_L + Y_H)/2$. For these agents the low regime history that $x_{t-1} \geq \omega_{M+1}$ is especially important: it limits the possible positions of the cluster to the left and thus imposes an upper bound on the value function. We begin with a simple result that highlights the issue.

PROPOSITION 5. For given values of $\alpha$ and $\beta$, there exists $\epsilon^*$ such that if $\epsilon < \epsilon^*$, in a low regime, $V_t(\omega_{M+2}) < \omega_{M+2}$.

*Proof.* For an agent with cost $\omega_{M+2}$, $W_t(c)$ is positive only if $x_t$ takes one of the two values $\omega_M$ or $\omega_{M+1}$. Since $\pi_t$ satisfies Property $\mathscr{H}$, the probability that $x$ takes one of these two values is vanishingly small when the number of grid points right of $c$ tends to become arbitrarily large; i.e., when $\epsilon$ becomes vanishingly small). More specifically, $W_t(\omega_{M+2}) = \pi_{M,t}(c)(2\epsilon) + \pi_{M+1,t}\epsilon$. From the updating equations (4), the probabilities $\pi_{M,t}(c)$ and $\pi_{M+1,t}(c)$ tend to zero like $\epsilon$. Hence $W_t(\omega_{M+2})$ tends to zero like $\epsilon^2$. For sufficiently small $\epsilon$, $W_t(\omega_{M+2}) < \epsilon < (1 - \beta)(\omega_{M+2} - c)$. The result follows from Lemma 2.

Proposition 4 shows that if there is a unique equilibrium the maximum cost of acting agents must be near $Y_L$. All agents know that the probability that $x < Y_L$ is small. But for agents with a cost $c$ near $Y_L$, and higher than $Y_L$, this is of special relevance because it puts a strong restriction on the probability that any agent in the cluster has a cost lower than $c$.

Consider now agents with a cost $c$ higher than $Y_L$ and not in the neighborhood of $Y_L$. They have a higher probability that some agents in the cluster have a cost lower than $c$, but they also need more of these agents to have a positive payoff of investing (as shown in Lemma 2). We now show that the history of a low regime also depresses the value function of these agents. Suppose that $c > Y_L$. Since $x_t$ is bounded below by $\omega_M \approx Y_L$, using the property $\mathscr{H}$ as in the proof of Lemma 3, one can find an upper bound of the mass of agents in the cluster with a cost smaller than $c$ by assuming that $x_t$ is uniformly distributed on the grid points between $\omega_M$ and $c$. Hence, $W_t(c) \leq \alpha(c - \omega_M)/2$, and if the grid is sufficiently fine, asymptotically, $W_t(c) \leq \alpha(c - Y_L)/2$. Using Lemma 2, this is a sufficient condition for $V_t(c) < c$ if Assumption 2 is satisfied. This argument is formalized in the Appendix to prove the next result.

LEMMA 4. If Assumption 2 holds and $\pi_t$ satisfies Property $\mathscr{H}$, there is $\epsilon^*$, which depends on the other parameters of the model, such that if $\epsilon < \epsilon^*$, then $V_t(c) < c$ for any $c$ with $\omega_{M+1} \leq c \leq (Y_L + Y_H)/2$.

The part of Proposition 4 which describes the low regime follows from Lemmata 3 and 4. A low regime ends in the first period for which $x_t = \omega_M$. At the end of that period the common knowledge distribution satisfies the Property $\mathscr{L}$ of the distribution with relatively low costs. A high regime begins in period $t + 1$. This regime is symmetrical to the previous one. The symmetry appears when we introduce the function $\tilde{W}_t(c)$ which measures the mass of agents in the cluster with a cost higher than $c$. Since the total mass of the cluster is $\alpha\sigma$, $\tilde{W}_t(c) = \alpha\sigma - W_t(c)$. The counterpart of Lemma 2 is that

$$V_t(c) > c \text{ if and only if } \tilde{W}_t(c) < (1 - \beta)(Y_H - c).$$

The arguments used for a low regime in Lemmata 3 and 4 can now be used for a high regime, mutatis mutandis, which concludes the proof of Proposition 4.

## IV. DISCUSSION

### *IV.A. The Impact of Extreme Events*

The existence of "extreme values" of the structural parameters is essential for the unique equilibrium with alternating regimes. These states are defined as the realizations of the cost parameter $x$ which are sufficiently low or sufficiently high such that there is a unique equilibrium with high or low activity, respectively, under perfect information. Assumption 1 ensures that the probability of these states is sufficiently important. This assumption was used in the formal analysis to generate Property $\mathscr{H}$ of the common knowledge distribution $\pi_t$ which puts higher probabilities on higher costs, in a low regime. However, this property is rather weak since it is satisfied by a uniform distribution to the right of $\omega_M$, whereas the actual bias of $\pi_t$ may be much stronger, (as shown in Figure III and in the asymptotic distribution of Proposition 6 in the Appendix). We can therefore conjecture that the uniqueness of the equilibrium holds when the sufficient condition of Assumption 1 is somewhat relaxed. Numerical simulations have verified this conjecture. For some parameters, the equilibrium is unique even if there is only one grid point $\omega_{M+1}$ greater than $Y_H$.

If the low equilibrium is impossible under perfect information (case (3) in Figure I), then $\omega_K < Y_H$, and the high value $Y_H$ always defines an equilibrium. Furthermore, if $\omega_1 < Y_L$, one can show (using the same techniques as in the previous section) that this is the only equilibrium. Thus, a small perturbation at one end of the range $x$ may have a very large impact: it shifts the economy from an equilibrium with alternating regimes where the economy spends significant time in low activity to a permanent level of high activity. Such a perturbation illustrates the impact of the extreme events. Note finally that if the range of $x \in [\omega_1, \omega_K]$, is contained in $(Y_L, Y_H - \sigma)$, then both $Y_L$ and $Y_H$ are equilibrium levels of activity in any period.

### *IV.B. Heterogeneity and History*

Each agent updates the common distribution $\pi_t$ which is established from history with the private information of his own cost. This updated distribution $\pi_t(c)$ shifts the common knowledge $\pi_t$ toward the value $c$. If agents are sufficiently heterogeneous, the observation of a cost $c$ does not convey much information about the cost of others and $\pi_t(c)$ is not very different from $\pi_t$. A sufficient degree of heterogeneity was provided by Assumption 2 in the formal analysis.

If the degree of heterogeneity is small (and the cost distribution is concentrated), then the observation of $c$ provides significant information on the cost of the others. Carlsson and Van Damme [1993] have remarked that when heterogeneity is vanishingly small, $\pi_t(c)$ may tend to a uniform distribution centered on $c$ (on a vanishing interval). We now show that in this case history does not matter for the agents with a cost which is not near $Y_L$ or $Y_H$, and that there is no longer a unique equilibrium.

Assume that $\beta$ and $\sigma$ are vanishingly small while the total mass of the cluster $\alpha\sigma$ is held invariant. Asymptotic values for $Y_L$ and $Y_H$ are 0 and $\alpha\sigma$.[7] We may assume that the economy is in a low regime, and we have to distinguish the cases of high and low costs, respectively.

Consider first an agent with cost $c \in [(Y_L + Y_H)/2 - \eta, (Y_L + Y_H)/2 + \eta]$, where $\eta$ is small but fixed as $\beta$ and $\sigma$ tend to 0. Because the distribution $\pi_t$ is relatively smooth, the information of his own cost dominates. When heterogeneity is vanishingly small, one can show that this agent believes that asymptotically half the agents have a cost lower than $c$ and the other half a cost higher than $c$:

$$\alpha\sigma/2 - \zeta < V(c) < \alpha\sigma/2 + \zeta, \tag{8}$$

where $\zeta$ is vanishingly small. From these inequalities there are two values $c_1$ and $c_2$ such that for small $\alpha$ and $\sigma$,

$$c_1 < (Y_L + Y_H)/2 < c_2, \quad \text{with} \quad V(c_1) > c_1, \quad \text{and} \quad V(c_2) < c_2.$$

Consider now an agent with an cost near $Y_L$, say, $c = \omega_{M+2}$ which is within $2\epsilon$ of $Y_L$. History in the low regime shows that the minimum cost in the cluster cannot be less than $Y_L - \epsilon$ (less than $\omega_M$ to be precise). Therefore, this agent cannot believe that nearly half the others have a cost lower than his. On the contrary, he believes that the probability that $x < c$ (i.e., $x = \omega_{M-1}$ or $x = \omega_M$), is very small. Proposition 5 showed that for sufficiently small $\epsilon$,

$$V(\omega_{M+2}) < \omega_{M+2}.$$

For the agents with a cost $c$ near $Y_L$, history dominates the information of $c$ no matter how small the heterogeneity.

From the previous inequalities, the criteria in Propositions 2 and 3 are not satisfied. A numerical example is illustrated by the graph of the value function in Figure V, which has three intersections with the diagonal,[8] say for the values $c_1^* < c_2^* < c_3^*$. The acting set $[-b, c_1^*]$ defines an equilibrium. However, in some

7. The value of $\epsilon$ must also tend to zero to have a sufficient number of grid points in the cluster. This can be done since Proposition 4 holds for small $\epsilon$. Furthermore, one could also assume a constant mass of agents with negative costs to keep the limit of $Y_L$ strictly positive.

8. In Figure V, $p = 0.1$, which is lower than the value in Figure V ($p = ⅓$). (Recent history is more informative if $p$ is small.)

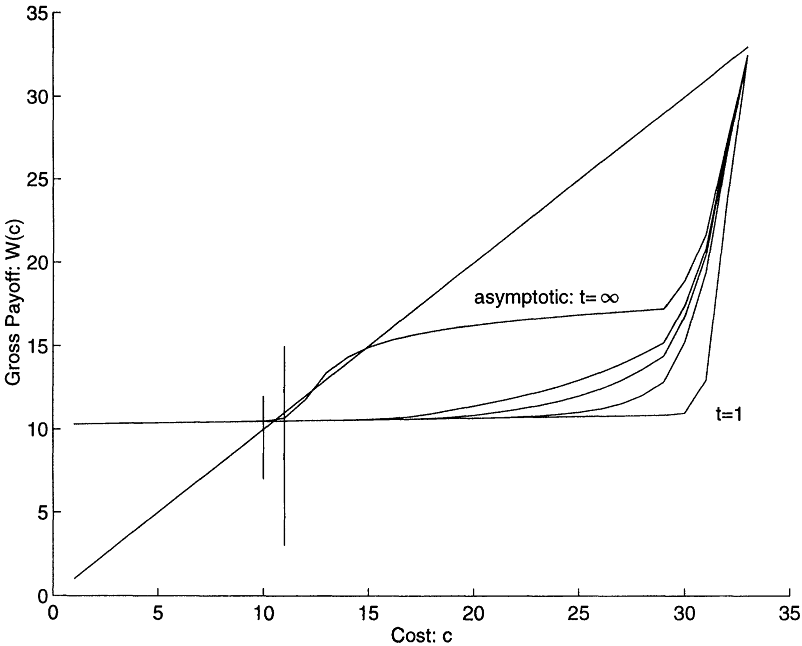


FIGURE V
The Case with Very Low Heterogeneity
Parameters: $\omega_M = 10 = Y_L - 0.5$, $\omega_N = 31 = Y_H - \sigma + 0.5$, $\omega_K = 31$, $\sigma = 2$, $\beta = 0.05$, $\alpha = 10.45, p = 0.1$.

cases, $[-b, c_3^*]$ may also define an equilibrium acting set.[9] Numerical simulations show that the existence of multiple equilibria requires a very low level of heterogeneity. For Figure V the density of costs is 40 times higher in the cluster than outside.[10]

The properties of the model can be compared with those of Carlsson and Van Damme [1993]. They analyze a similar "stag hunt" game in which heterogeneity is vanishingly small, and which is played only once. They were the first to emphasize how extreme events can be a determining factor for the regions of iteratively dominant strategies. Adapting their result to the present framework, for vanishingly small heterogeneity and if the

9. For the case of Figure V the results of this paper are not sufficient to establish whether $[-b, c_3^*]$ defines an equilibrium action set. However, by continuity, one may safely conjecture that when the structural parameters are adjusted such that $V$ shifts upward and $c_2^*$ tends to $c_1^*$, $c_3^*$ indeed defines an equilibrium.

10. As an example of robustness when Assumptions 1 and 2 are relaxed, the equilibrium is unique when $K = N$, $\beta = 0.2$, and $\alpha = 8$.

game were played only once, the positive probabilities of each of the two types of extreme events (with low and high costs, respectively), would induce agents with a cost lower than $(Y_L + Y_H)/2 - \zeta$ to act (by iterative dominance), and those with a cost higher than $(Y_L + Y_H)/2 + \zeta$ not to act (with $\zeta$ vanishingly small). This argument applies here when $c > (Y_L + Y_H)/2 + \zeta$. But it fails for $c < (Y_L + Y_H)/2 - \zeta$ because history overrules the effect of extreme events for the agents near the critical value $Y_L$. In summary, Carlsson and Van Damme showed that when a coordination game is played once, for most payoff values one of the Nash equilibria under perfect information is "selected" in the game with small heterogeneity when agents play strategies that are iteratively dominant. This paper indicates that this setting is inappropriate when the game is played repeatedly, by different agents in each period, and structural parameters change by small amounts between periods.

### *IV.C. Observation Lags and Random Walk Drift*

The model assumes that an observation takes place after each move of $x$, and that the grid on which $x$ moves is sufficiently fine. It can be viewed as an approximation of a model with continuous time. However, observations of macroeconomic data are often made with significant lags. Such lags can be introduced here by assuming a finite number of steps for the random moves of $x$ (as specified in (2)) between periods. Numerical simulations[11] provide some support for the robustness of a unique equilibrium with random switches. The equality of the probabilities for increases and decreases of $x$ can also be relaxed, somewhat. By choosing appropriate values for these probabilities and for $\omega_M$, $\omega_N$, and $\omega_K$, it is possible to reproduce a large set of transition probabilities between the two regimes (including values that are similar to those of Hamilton [1989]).

### *IV.D. The Shape of Agents' Distribution*

The rectangular density function is now replaced by a smooth function: the population is divided into two groups as in the previous sections; the first has a uniform density of cost $\beta$ on the interval $[-b, B]$; the second group is the cluster and is distributed normally $N(x_t,\sigma)$, where $\sigma$ is constant (publicly known), and the mean of the cluster, $x_t$, follows a random walk as in equations (2).

11. See Chamley [1996].

The mass of acting agents is equal to $Y_t$, as in the standard model, and agents observe at the end of period $t$ the variable $Z_t$ which is defined by

$$Z_t = Y_t + \eta_t,$$

where $\eta_t$ are i.i.d. random variables, normally distributed $N(0,\omega)$, and $\omega$ is publicly known.[12] The variable $\eta$ is interpreted as an observation noise or as the activity of "noise traders" who act independently of the level of aggregate activity. The observation noise is both plausible and necessary, given the parsimonious parameterization of the structure of the economy. With no noise one observation would reveal the complete structure of all costs in the economy, and that would be a nongeneric property. As in the previous sections, the payoff of an agent $c$ is equal to the expected value $E[Y_t] - c$.

Suppose that the acting set is the interval $[-b, c^*]$. The signal of aggregate activity is

$$Z = \beta(b + c^*) + F(c^*; x) + \eta, \tag{6}$$

where $F(c^*; x)$ is the c.d.f. at the point $c^*$ of the normal distribution $N(x,\sigma)$. The signal extraction problem is to learn about $x$ from $Z$, which depends on $x$ through the function $F(c^*; x)$. When $|c^* - x|$ is large, $F(c^*; x)$ does not depend much on $x$ and tends to 0 or 1. In that case, the noise $\eta$ dwarfs the impact of $x$ on $F(c^*; x)$, and the observation of $Z$ conveys little information on $x$. Learning is significant only of $|c^* - x|$ is relatively small, i.e., when the density function $f(c^*; x)$ is sufficiently high.[13] But the strength of strategic complementarity is positively related to $f(c^*; x)$ (which is identical to the slope of the reaction function under perfect information). *Learning and strategic complementarity are positively related.* As in the previous model with rectangular densities, agents learn a significant amount of information only when the density of agents near a critical point is sufficiently large to push the economy to the other regime.

The model cannot be analyzed analytically, but numerical simulations[14] have shown that the main features of the standard model still hold. One verifies that for most realizations of $x_t$ there

12. Negative values of $Z_t$ have a very small probability and are neglected as in all linear econometric models with normal distributions.
13. Such a model is also used in the analysis of delays and multiple equilibria [Chamley 1997].
14. See Chamley [1996] for the details of the solution and the results.

is a unique equilibrium. As anticipated, the amount of information conveyed by the aggregate activity is small when the mean of the population is relatively far from the critical regions where a switch of regimes might occur, and it increases dramatically when such a transition occurs.

The model does not have a unique equilibrium for all realizations of $x_t$, however. Consider the case of low regime, and suppose that the true value of $x$ is in the range for which there are two equilibria under perfect information. A very large positive shock $\eta$ might occur (with a very small probability). It could fool a significant mass of agents to act. In this case, the amount of information released by the observation of aggregate activity would be large and could generate multiple equilibria for the following period. The numerical simulations indicate that, subject to the standard parameter restrictions, such events while possible, seem rather rare.

The purpose of the model is not to resolve the problem of multiple equilibria in all cases. It is to resolve this problem for the kind of uncertainty and heterogeneity that are relevant in macroeconomics or in other contexts of social behavior.

## V. Applications

### V.A. Policy

When the economy is in a low regime, a fiscal policy that pays a subsidy for action increases the acting set and thus the probability of a transition to a high regime. In the present model, the effect of such a stimulus is to shift the cluster of costs to the left. The model shows that a subsidy (of moderate size) has little effect if individuals expect the average cost in the cluster to be relatively high.

In the context of the business cycle with payoff externalities, policy is not very effective in the first phase of a recession. For a given amount of subsidy, it may be advisable to wait until the common knowledge probability of an upswing reaches a higher level. A policy is more effective when people are more optimistic about the economy. The model shows why there is more in this argument than an apparent triviality.

The impact of policy depends also on the results of past policies in a way that cannot be represented by the ad hoc models of equilibrium selection which were discussed in the introduction.

Suppose that a subsidy is implemented in period $t$. The value of the maximum cost of acting agents $c_t^*$ increases. If the policy fails and does not generate a recovery, agents learn that $x_t > c_t^*$ and are more pessimistic in period $t + 1$ than they would have been with no policy in period $t$. Hence, a repetition of the subsidy in period $t + 1$ is much less effective ex ante than the subsidy in period $t$.

### *V.B. Social Changes and Revolutions*

Kuran [1995] asked why sudden changes of opinions or revolutions that were not anticipated with high probability seem anything but surprising in hindsight. The gap between the ex ante and the ex post view is especially striking when no important exogenous event occurs (e.g., the French revolution, the fall of communist regimes). These social changes depend essentially on the distribution of individuals' payoffs, on which each one has only partial information. As Kuran argued, "historians have systematically overestimated what revolutionary actors could have known." If a revolution were to be fully anticipated, it would probably run a different course. Louis the XVI entered in his diary "nothing" on July 14, 1789. Before a social change, individuals who favor the change do not have perfect information on the preference of others ex ante, but they are surprised to find themselves in agreement with so many ex post and this common view in hindsight creates a sense of determinism.

Kuran's model is static and isomorphic to other models with an ad hoc rule of equilibrium selection by proximity. But the analysis of changes and surprises requires a dynamic approach with an explicit formation of expectations. Following Kuran [1988, 1995], suppose that individuals have to decide between two "expressed opinions" or "attitudes" as revealed by some behavior: action 1 supports a given political regime, while action 0 does not (or supports a revolution). Each individual is characterized by a preference variable $c$ which is distributed on the interval [0,1], with a cumulative distribution function $F(c)$. The preference for the regime increases with $c$. There is a continuum of individuals with a total mass $\Omega$. In any period, individuals have to choose an action $s$ which is equal to either 0 or 1. The payoff of action $s$ for an individual with parameter $c$ is (i) a decreasing function of the "distance" between his action and his preference parameter, (ii) an increasing function of the mass of individuals who choose the same action. Kuran interpreted the externality effect by an individual's taste for conformity. Strategic complementarities may also arise because the probability of the change of regime depends on the number of individuals expressing an opinion, or taking an active part in a revolution.

Denote by $Y$ the mass of individuals who chose action 0 (the revolution) in a given period, (and by $\Omega - Y$ the mass of individuals who choose action 1). From the previous discussion, the payoff function of an individual who take action $s$ is of the form,

$$w(s, Y, c) = \begin{cases} Y - c, & \text{if } s = 0, \\ \Omega - Y - (1 - c), & \text{if } s = 1. \end{cases}$$

The difference $w(0, Y, c) - w(1, Y, c)$ is a function $u$ such that

$$u(c) = 2(Y - c) - \Omega + 1, \qquad \text{with} \qquad c \in [0, 1],$$

which is just a linear transformation of the payoff function $Y - c$ posited in the present paper. Kuran's model is thus a special case of the canonical model with strategic complementarities. For a suitable distribution of individual preferences, it has multiple equilibria under perfect information. A regime is assumed to prevail as long as the structure of preferences allows it. This structure may evolve such that the regime is no longer a feasible equilibrium, and society jumps to the other equilibrium regime. The concepts of surprises and expectations, which cannot be analyzed in a static approach, have a prominent place in the dynamic models of this paper.[15]

Until the very end of the old regime, the common knowledge, which is also the probability assessment of an outsider with no individual preference, is that a large fraction of the population supports the old regime (as indicated by the bias of $\pi_t$ toward higher costs), whereas the actual distribution can support a revolution. When the regime changes, expectations change in two ways: first, the perceived distribution of preferences shifts abruptly

15. The sudden change of expectations and the "wisdom-after-the-fact" effect in this paper is also found in Caplin and Leahy [1994]. There are important differences, however. As emphasized in subsection IV.D following equation (6), the outburst of information is related to strategic complementarity. There is no strategic complementarity in Caplin and Leahy. In this model, the outburst of information occurs when a subgroup of agents receives a string of identical messages, say negative, over a number of periods sufficiently large to prompt them to change their action. Because their number is infinite, their behavior provides perfect information. Other agents may or may not follow after learning. In an ergodic version of their model (comparable to the setting here), the outburst of information would occur periodically even if the level of aggregate activity does not switch to a new regime. Sudden bursts of information and jumps of aggregate activity appear in Zeira [1994] with Bayesian learning and strategic substitutabilities.

toward the new regime; second, the precision of this perception is much more accurate. The high confidence in the information immediately after the revolution may provide all individuals with a sense that the revolution was deterministic.

Further work may apply the Bayesian approach to an analysis of the policy for an authority who attempts to maintain the old regime (or for the revolutionaries). We have already seen in the previous section how the removal of a penalty for "action 0" may increase the probability of a switch to another equilibrium: as noted by de Tocqueville [1856], regimes crumble not when they are at their most repressive, but when this state of repression becomes partially lifted.[16]

### V.C. Further Issues

This paper began with a standard model of strategic complementarities with multiple equilibria under perfect information. The introduction of imperfect information on individual parameters and individual rational behavior reduced the set of possible equilibria to a unique one that exhibits random transitions between phases of high and low levels of activity. The properties of the model are not valid for all possible parameter values, and obviously cannot be, as we have seen. However, they are robust for the types of uncertainty and heterogeneity that are plausible in the contexts of macroeconomics or social regimes.

Further research could investigate three issues. The first is to generalize the structure of preferences or technology. The second is the introduction of delays. When agents can delay an irreversible action, they may use this opportunity to wait for more information. (Chamley and Gale [1994] focus exclusively on informational externalities.) In the present model with no delay, the states of high and low activity are completely symmetrical. Delays may stretch regimes of low activity while shortening those with high activity. They may thus introduce differences between the phases of aggregate activity, with possible implications for the time series properties of endogenous cycles and the impacts of policy. Finally, the dynamic approach to stag hunt games, which is proposed here, may find other applications in economics and social behavior.

16. Tocqueville had probably more than one mechanism in mind: before the revolution, "*le mal est devenu moindre il est vrai, mais la sensibilité est plus vive*" (Chapter 4).

## APPENDIX

### *Proof of Proposition 2*

Define $\hat{c}^k = \sup\{c|[-b, c) \subset A^k\}$. The sequence $\{c^k\}$ is increasing. Since it is bounded, it converges to some $\bar{c}$. The inequality $\bar{c} \geq c^*$ is now proved by contradiction. Suppose that $\bar{c} < c^*$:

$$E[\mu_x(-b, c^k)|\hat{c}] > E[\mu_x(-b, \hat{c})|\hat{c}] - (\alpha + \beta)(\hat{c} - c^k)$$

$$= V(\hat{c}) - (\alpha + \beta)(\hat{c} - c^k).$$

The function $V$ is continuous except on some grid points.

Consider the compact interval $[-b,(\bar{c} + c^*)/2]$, which is strictly included in $[-b,c^*]$ and strictly contains $[-b,\bar{c}]$. There exists $\delta > 0$ such that $V(c) > c + \delta$ for all values $c \in [-b, (\bar{c} + c^*)/2]$. (This is proved by contradiction as it would be if $V$ were continuous.)

Hence, for any $c$ such that $c \leq (\bar{c} + c^*)/2$,

$$E[\mu_x(-b, c^k)|c] > c - (\alpha + \beta)(c - c^k) + \delta,$$

and for any $c$ such that $c^k < c < \min(c^k + (\alpha + \beta)/\delta,\ \bar{c} + (c^* - \bar{c})/2)$,

$$E[\mu_x(-b, c^k)|c] > c.$$

Choosing $c^k > \bar{c} - (\alpha + \beta)/\delta$ shows that $c^{k+1} > \bar{c}$, which contradicts the definition of $\bar{c}$.

The second part of the result is proved by the same method.

### *The Asymptotic Limit of $\pi_t$ in a Low Regime*

PROPOSITION 6. Call $\hat{\pi}_{k,t}$ the value of $\pi_{k,t}$ conditional on the observation $x_t > Y_L$. If the economy is in a low regime in periods $\tau \leq t$, and if $x_t > Y_L$, then

$$\mathrm{Lim}_{t\to\infty}\hat{\pi}_{k,t}$$

$$= \begin{cases} 0, & \text{for } \omega_k < Y_L, \\ 2\tan\left(\dfrac{\omega_0}{2}\right)\sin(\omega_0(k - M)), & \text{with } \omega_0 = \dfrac{\pi}{2(K - M) + 1} \end{cases}$$

*Proof of Proposition 6*

Suppose that $x_t \geq \omega_{M+1}$. At the end of period $t$, agents update the assessment $\pi_t$ to $\hat{\pi}_t$ which is defined by

$$\text{(A1)} \qquad \hat{\pi}_{k,t} = \begin{cases} 0, & \text{for } 1 \leq k \leq M, \\ \dfrac{\pi_{k,t}}{1 - \pi_{M,t}}, & \text{for } M + 1 \leq k \leq K, \end{cases}$$

For period $t + 1$, the updating formulas are

$$\text{(A2)} \quad \pi_{k,\,t+1} = \begin{cases} 0, & \text{for } 1 \leq k \leq M - 1, \\ p\hat{\pi}_{k-1,t} + (1 - 2p)\hat{\pi}_{k,t} \\ \quad + p\hat{\pi}_{k+1,t}, & \text{for } M \leq k \leq K - 1, \\ p\hat{\pi}_{K-1,t} + (1 - p)\hat{\pi}_{K,t}, & \text{for } k = K. \end{cases}$$

Given $\hat{\pi}_t$, the application of equation (A2) and then of equation (A1) yields $\hat{\pi}_{t+1}$. If the sequence $\hat{\pi}_t$ has a limit $\hat{\pi}$, it is a fixed point of the updating equations (A1) and (A2). Define the values $z_k$ such that $\hat{\pi}_{M+k} = z_k$, and $R = K - M$. We have $z_k = 0$ for $-M + 1 \leq k < 0$, and

$$\text{(A3)} \qquad \begin{cases} z_0 = 0, z_1 = a, \\ z_{k+1} - (2 - a)z_k + z_{k-1} = 0 \text{ for } 1 \leq k \leq R - 2, \\ z_R = z_{R-1}/(1 - a). \end{cases}$$

The last equation $z_R = z_{R-1}/(1 - a)$ can be replaced by the introduction of a variable $z_{R+1}$ with the condition $z_{R+1} = z_R$, and extending the previous equation for $k = R - 1$.

Since the polynomial $z^2 - (2 - a)z + 1$ has complex roots of modulus one, $z_k$ is of the form $z_k = C \sin(\omega k + \phi)$, for some parameters, $C$, $\omega$, $\phi$. $\phi = 0$ because $z_0 = 0$, and

$$\text{(A4)} \qquad z_k = C \sin(\omega k), \text{ for } 0 \leq k \leq R + 1.$$

Using the equations $z_2 = (2 - a)z_1$, $\alpha$ and $z_R = z_{R+1}$, respectively, we have

$$\text{(A5)} \quad \cos(\omega) = 1 - a/2,\ C \sin(\omega) = a,\ \sin(\omega R) = \sin(\omega R + \omega).$$

By definition of the probabilities, $\sin(\omega k) \geq 0$ for all $k$ such that $0 \leq k \leq R + 1$. From these inequalities and $\sin(\omega R) = \sin(\omega R + \omega)$, $\pi/2 - \omega R = \omega R + \omega - \pi/2$, or

$$\text{(A6)} \qquad \omega = \pi/(2R + 1).$$

This equation defines a unique value for $\omega$. The values of $a$ and $C$ are then determined in equation (A5): $C \sin(\omega) = a = 2(1 - \cos(\omega)) = 4 \sin^2(\omega/2)$. Hence,

$$C = 2 \tan(\omega/2), \text{ and } a = 2(1 - \cos(\omega)).$$

Therefore, if a limit of $\hat{\pi}_t$ exists, it is unique.

Since $\hat{\pi}_t$ is a probability distribution, the transformation from $\hat{\pi}_t$ to $\hat{\pi}_{t+1}$ is defined on a compact set. By extracting subsequences of the sequences $\hat{\pi}_{k,t}$ for each $k$ ($M \leq k \leq K$), which all belong to the compact set, each of them converges to $\pi_k^*$, which proves the existence part of the proposition. The rest of the proposition follows immediately from the equations (A5) and (A6).

*Proof of Lemma 3*

Define $J(y)$ as the index of the grid point nearest to $y$ and smaller than $y$:

$$y - \epsilon \leq \omega_{J(y)} < y.$$

As in the description of the text, we distinguish two cases.

A. Consider first an agent with cost $c > \omega_N + \sigma/2$. Define the function,

$$\psi(\zeta; c) = \min(\max(\alpha(c - \zeta), 0), \alpha\sigma).$$

It is the mass of the cluster left of $c$ when $x = \zeta$. Its graph is represented in Figure VI:

$$W_t(c) = \sum_k \pi_{k,t}(c)\psi(\omega_k; c).$$

In this expression, pair the terms with the same "distance" from $\omega_N$:

$$W_t(c) = \sum_{\substack{N+i \leq K \\ N-i \geq M}} \pi_{N+i,t}(c)\psi(\omega_{N+i}; c) + \pi_{N-i,t}(c)\psi(\omega_{N-i}; c).$$

From Property $\mathscr{H}$, for all values of $i \geq 1$ in the previous sum, $\pi_{N-i,t} \leq \pi_{N+i,t}$. Using $c > \omega_N + \sigma/2$, and the expressions (4) in the text for $\pi_{j,t}(c)$, the same "bias" property holds for the distribution $\pi_t(c)$: $\pi_{N-i,t}(c) \leq \pi_{N+i,t}(c)$.

Using the definition of $\psi$ and $c > \omega_N + \sigma/2$, $|\omega_N - (\omega_{J(c)} - \sigma)| \leq |\omega_N - \omega_{J(c)}|$ for $c \geq \omega_N + \sigma/2$: in Figure VI, the point $\omega_N$ is left of the middle of the segment $(\omega_{J(c)} - \sigma, \omega_{J(c)})$. Hence,

$$[\psi(\omega_{N+i}; c) + \psi(\omega_{N-i}; c)]/2 \leq \psi(\omega_N; c).$$

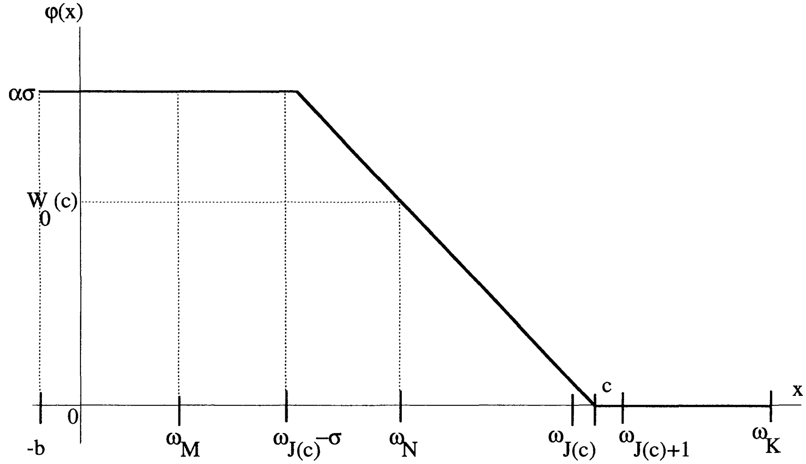


FIGURE VI
$\psi(x;c)$ as a Function of $x$ ($c$ Fixed)

Using this inequality and $\pi_{N-i,\,t}(c) \le \pi_{N+i,t}(c)$,

$$\pi_{N+i,t}(c)\psi(\omega_{N+i};c) + \pi_{N-i,t}(c)\psi(\omega_{N-i};c) \le (\pi_{N+i,t}(c) + \pi_{N-i,t}(c))\psi(\omega_N;c).$$

Summing over all possible values of $i$ in the previous expression of $W_t(c)$, we find that $W_t(c) \le \psi(\omega_N;\, c)$. Since $V_t(c) = W_t(c) + \beta(c + b)$, and by definition of $\omega_N$, $\psi(\omega_n;\, c) + \beta(c + b) < c$, $V_t(c) < c$.

B. Consider now an agent with $c \le \omega_N + \sigma/2$. Formalizing the argument in the text as in Part A above, and using Assumption 1, for any $t \ge 0$ in a low regime, $W_t(c) \le \alpha\sigma/2$. Because of Lemma 2, $V_t(c) < c$ if $W_t(c) < (1 - \beta)(c - Y_L)$, for which a sufficient condition is $\alpha\sigma/2 < (1 - \beta)(c - Y_L)$, or

$$c > Y_L + \alpha\sigma/2(1 - \beta) = (Y_L + Y_H)/2.$$

*Proof of Lemma 4*

For an agent with cost $c$, the expected mass of the cluster left of $c$ is equal to

$$W_t(c) = \sum_{k=M}^{k=J(c)} \pi_{k,t}(c)\alpha(c - \omega k) = S,$$

because $\pi_{k,\,t}(c) = 0$ for $k < M$, (Property $\mathcal{H}$). We now find an upper bound of this expression using the rightwards bias of $\pi_t$ in Property $\mathcal{H}$.

Let us first prove that $c < \omega_M + \sigma$. Since $c \leq (Y_H + Y_L)/2$ and $\omega_M = Y_L - \epsilon/2$, then

$$c - \omega_M \leq \frac{Y_H + Y_L}{2} - Y_L + \frac{\epsilon}{2} = \frac{Y_H - Y_L}{2} + \frac{\epsilon}{2} = \frac{\alpha\sigma}{2(1 - \beta)} + \frac{\epsilon}{2}.$$

Since $\alpha < 2(1 - \beta)$ by Assumption 2, there exists $\epsilon_1^*$ such that if $\epsilon < \epsilon_1^*$, $c < \omega_M + \sigma$. We assume that $\epsilon$ satisfies this condition.

Since $c < \omega_M + \sigma$, using the updating equation (4) in the text, for $M \leq k \leq J(c)$, $\pi_{k,t}(c) = \lambda\pi_{k,t}$ where $\lambda$ is a positive number. Hence the items (ii) and (iii) of Property $\mathscr{H}$ apply to $\pi_{k,t}(c)$ if $M \leq k \leq J(c)$. We distinguish two cases.

A. Suppose first that $c \geq \omega_K$ which is the highest point on the grid of values of $x$. By pairing the terms equidistant from $\omega_N$ and using Property $\mathscr{H}$, as in the proof of Lemma 3, an upper bound for $S$ is obtained if we replace the distribution $\pi_t(c)$ by a uniform distribution for all the grid points $x$ such that $\omega_M \leq x \leq \omega_K$. Therefore,

$$S \leq \alpha[c - (\omega_M + \omega_K)/2].$$

To prove that $V_t(c) < c$, using Lemma 2 and the definition of $\omega_N$, it is sufficient to prove that

$$\alpha[c - (\omega_M + \omega_K)/2] < (1 - \beta)(c - Y_L).$$

Since $\alpha > 1 - \beta$, and $c < \omega_M + \sigma$, it is sufficient to prove that the inequality holds when $c$ is replaced by $\omega_M + \sigma$. It is therefore sufficient to show that

$$\alpha\frac{\omega_M - \omega_K}{2} + \alpha\frac{\sigma}{2} < (1 - \beta)\left(\sigma - \frac{\epsilon}{2}\right),$$

which is equivalent to $\omega_K > \omega_M + \sigma(1 - 2(1 - \beta)/\alpha) + ((1 - \beta)/\alpha)\epsilon$. By Assumption 2, $\alpha < 2(1 - \beta)$. Since $\omega_K \geq \omega_N > \omega_M$, there is $\epsilon_2^*$ (which is not particularly small) such that this inequality is satisfied when $\epsilon < \epsilon_2^*$.

B. Suppose now that $c \leq \omega_K$. Using the previously demonstrated identity $\pi_{k,t}(c) = \lambda\pi_{k,t}$ for some $\lambda > 0$, Property $\mathscr{H}$ for $\pi_t$, and the same argument as before, an upper bound for $W(c)$ is obtained if we replace the distribution $\pi_t(c)$ by a uniform distribution for all the grid points left of $c$. Hence,

$$W_t(c) \leq \alpha(c - (\omega_M + \omega_{J(c)})/2).$$

Using the definitions of $\omega_M$ and $\omega_{J(c)}$, we have

$$W_t(c) \leq \alpha\left(c - \frac{Y_L + c}{2} + \epsilon\right) = \alpha\frac{c - Y_L}{2} + \alpha\epsilon.$$

From Assumption 2, there exists $\theta$ with $0 < \theta < 1$ such that $\alpha = 2(1 - \beta)(1 - \theta)$. Hence, $W_t(c) \leq (1 - \beta)(1 - \theta)(c - Y_L + \epsilon/2)$. Using Lemma 2, a sufficient condition for $V_t(c) < c$ is $(1/\theta - 1)\epsilon/2 < c - Y_L$, which is equivalent to $c > \omega_M + \epsilon/(2\theta)$. Define $\ell$ as the smallest integer such that $\ell > 1/(2\theta)$ The inequality is satisfied for $c \geq \omega_{M+\ell} \geq \omega_M + \epsilon/(2\theta)$. If $\ell > 1$, by a straightforward extension of Proposition 5, (using the property that $\ell$ is independent of $\epsilon$ and finite), there is $\epsilon_3^*$ such that if $\epsilon < \epsilon_3^*$, $V_t(c) < c$ for $c \in [\omega_{M+1}, \omega_\ell)$.

BOSTON UNIVERSITY AND EHESS

## REFERENCES

Bryant, John, "The Paradox of Thrift, Liquidity Preference and Animal Spirits," *Econometrica,* LXI (1987), 1231–1235.
Caplin, Andrew, and John Leahy, "Business as Usual, Market Crashes and Wisdom after the Fact," *American Economic Review,* LXXXIV (1994), 547–564.
Carlsson, Hans, and Eric Van Damme, "Global Payoff Uncertainty and Risk Dominance," *Econometrica,* LXI (1993), 989–1018.
Chamley, Christophe, "Coordination of Heterogeneous Agents in a Unique Equilibrium with Random Regime Switches," DELTA Document No. 96-17, 1996.
____, "Social Learning, Delays, and Multiple Equilibria," mimeo, 1997.
Chamley, Christophe, and Douglas Gale, "Information Revelation and Strategic Delay in a Model of Investment," *Econometrica,* LXII (1994), 1065–1085.
de Tocqueville, Alexis (1856), *L'Ancien Régime et la Révolution,* Robert Laffont, ed. (1986).
Cooper, Russell, "Equilibrium Selection in Imperfectly Competitive Economies with Multiple Equilibria," *Economic Journal,* CIV (1994), 1106–1122.
Cooper, Russell, and Andrew John, "Coordinating Coordination Failures in Keynesian Models," *Quarterly Journal of Economics,* CIII (1988), 441–463.
Diamond, Peter, "Aggregate Demand Management in Search Equilibrium," *Journal of Political Economy,* XC (1982), 881–894.
Diebold, Francis, Georges Rudebusch, and David Sichel, "Further Evidence on Business Cycle Duration Dependence," in James Stock and Mark Watson, eds., *Business Cycles, Indicators and Forecasting* (Chicago: The University of Chicago Press, 1993), pp. 255–280.
Filardo, Andrew, "Business-Cycle Phases and their Transitional Dynamics," *Journal of Business & Economic Statistics,* XII (1994), 299–308.
Fudenberg, Drew, and Jean Tirole, *Game Theory* (Cambridge, MA: MIT Press, 1991).
Goodwin, Richard, "The Non-Linear Accelerator and the Persistence of Business Cycles," *Econometrica,* XIX (1951), 1–17.
Hamilton, James, "A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle," *Econometrica,* LVII (1989), 357–384.
Kiyotaki, Nobuhiro, "Multiple Expectational Equilibria under Monopolistic Competition," *Quarterly Journal of Economics,* CIII (1988), 695–713.

Kuran, Timur, "Preference Falsification, Policy Continuity and Collective Conservatism," *Economic Journal,* XCVII (1987), 642–665.

——, *Legacies of Living a Lie* (Cambridge, MA: Harvard University Press, 1995).

Morris, Stephen, "Cooperation and Timing," University of Pennsylvania, mimeo, 1995.

Murphy, Kevin, Andrei Shleifer, and Robert Vishny "Industralization and the Big Push," *Journal of Political Economy,* XCVII (1989), 1003–1023.

Zeira, Joseph, "Information Cycles," *Review of Economic Studies,* LXI (1994), 31–44.