MIT LIBRARIES DUPL
3 9080 02617 8241

Digitized by the Internet Archive
in 2011 with funding from
Boston Library Consortium Member Libraries

http://www.archive.org/details/coordinationpoli00ange2

30²
31
415
.03-
.8
003

Massachusetts Institute of Technology
Department of Economics
Working Paper Series

COORDINATION AND POLICY TRAPS

George-Marios Angeletos
Christian Hellwig
Alessandro Pavan

Working Paper 03-18

May 2003
Rev. September 2003

Room E52-251
50 Memorial Drive
Cambridge, MA 02142

This paper can be downloaded without charge from the
Social Science Research Network Paper Collection at
http://ssrn.com/abstract=410100

MASSACHUSETTS INSTITUTE
OF TECHNOLOGY
JAN 2 1 2004
LIBRARIES

Coordination and Policy Traps*

George-Marios Angeletos
MIT and NBER

Christian Hellwig
UCLA

Alessandro Pavan
Northwestern University

This draft: September 2003

Abstract

This paper examines the ability of a policy maker to fashion equilibrium outcomes in an environment where market participants play a coordination game with heterogeneous information. We consider a simple model of regime change that embeds many applications examined in the literature. In equilibrium, the policy maker is willing to take a costly policy action only for moderate fundamentals. Market participants can use this information to coordinate on different responses to the same policy choice, thus inducing *policy traps*, where the optimal policy and the resulting regime outcome are dictated by self-fulfilling market expectations. Despite equilibrium multiplicity, *robust predictions* can be made. The probability of regime change is monotonic in the fundamentals, the policy maker intervenes only in a region of intermediate fundamentals, and this region shrinks as the information in the market becomes precise.

Key Words: strategic complementarities, global games, signaling, regime change, market expectations, policy.

JEL Classification Numbers: C72, D70, D82, D84, E52, E61, F31.

*For encouragement and helpful comments, we are thankful to Daron Acemoglu, Andy Atkeson, Philippe Bacchetta, Gadi Barlevy, Marco Bassetto, Olivier Blanchard, Ricardo Caballero, V.V. Chari, Eddie Dekel, Glenn Ellison, Paul Heidhues, Patrick Kehoe, Robert Lucas Jr., Narayana Kocherlakota, Kiminori Matsuyama, Stephen Morris, Nancy Stokey, Jean Tirole, Muhamet Yildiz, Ivan Werning, and seminar participants at AUEB, Berkeley, Carnegie-Mellon, Harvard, Iowa, Lausanne, LSE, Mannheim, MIT, Northwestern, Stanford, Stony Brook, Toulouse, UCLA, the Minneapolis Fed, the 2002 SED annual meeting, the 2002 UPF workshop on coordination games, the 2003 CEPR ESSIM, and the 2003 SITE workshop. *Email addresses:* angelet@mit.edu, chris@econ.ucla.edu, alepavan@northwestern.edu.

As Mr. Greenspan prepares to give a critical new assessment of the monetary policy outlook in testimony to Congress on Wednesday, the central bank faces a difficult choice in grappling with the economic slowdown. If it heeds the clamour from much of Wall Street and cuts rates now ... it risks being seen as panicky, jeopardizing its reputation for policy-making competence. But if it waits until March 20, it risks allowing the economy to develop even more powerful downward momentum in what could prove a crucial three weeks. (Financial Times, February 27, 2001)

1 Introduction

Economic news anxiously concentrate on the information that different policy choices convey about the competence and intensions of the policy maker, how markets may interpret and react to different policy measures, and whether government intervention can calm down animal spirits and ease markets to coordinate on desirable courses of actions. This paper investigates the ability of a policy maker to influence market expectations and control equilibrium outcomes in environments where market participants play a coordination game with heterogeneous information.

A large number of economic interactions are characterized by strategic complementarities, can be modeled as coordination games, and often exhibit multiple equilibria sustained by self-fulfilling expectations. Prominent examples include bank runs (Diamond and Dybvig, 1983), currency crises (Obstfeld, 1986, 1996; Velasco, 1996), debt crises (Calvo, 1988; Cole and Kehoe, 1996), financial crashes (Freixas and Rochet, 1997; Chari and Kehoe, 2003a,b), and regime switches (Chamley, 1999). Strategic complementarities also arise in economies with production externalities (Bryant, 1983; Benhabib and Farmer, 1994), network externalities (Katz and Shapiro, 1985, 1986; Farell and Saloner, 1985), imperfect market competition (Milgrom and Roberts, 1990; Vives 1999), Keynesian frictions (Cooper and John, 1988; Kiyotaki, 1988), thick-market externalities (Diamond, 1982; Murphy, Shleifer and Vishny, 1989), restricted market participation (Azariadis, 1981), and incomplete financial markets (Bernanke and Gertler, 1989; Kiyotaki and Moore, 1997; Angeletos, 2003).[1] Other examples include lobbying, political reforms, revolutions, riots, and social change (Kuran, 1987; Atkeson, 2000; Battaglini and Benabou, 2003).

Building on the *global-games* work of Carlsson and Van Damme (1993), Morris and Shin (1998, 2000, 2001) have recently argued that equilibrium multiplicity in such coordination environments

[1] An excellent overview of coordination games in macroeconomics is in Cooper (1999).

is often "the unintended consequence" of assuming common knowledge of the payoff structure of the game (the "fundamentals"), which implies that agents can perfectly forecast each others' beliefs and actions in equilibrium. When instead different agents have different private information about the underlying fundamentals, perfect coordination on multiple courses of action is no longer possible.[2] This argument has been elegantly illustrated by Morris and Shin (1998) in the context of self-fulfilling currency crises. When speculators receive noisy idiosyncratic signals about the willingness and ability of the monetary authority to defend the currency (the fundamentals), a unique equilibrium survives, in which devaluation occurs if and only if the fundamentals fall below a critical state. This unique threshold in turn depends on all policy variables affecting the payoffs of the speculators. For example, raising domestic interest rates, imposing capital controls, or otherwise increasing the cost of attacking the currency, reduces the set of fundamentals for which devaluation occurs. Morris and Shin hence argue that, "in contrast to multiple equilibrium models, [their] model allows analysis of policy proposals directed at curtailing currency attacks."

However, policy choices also convey information. For example, when the fundamentals are so weak that the collapse of the currency is inevitable, there is no point in adopting costly defense measures. Similarly, when the fundamentals are so strong that the size of the attack is minuscule and the currency faces no threat, there is no need to act. Therefore, whenever the policy maker intervenes, the market can infer that the fundamentals are neither too weak nor too strong. This information in turn may permit coordination on multiple courses of action in the market, thus interfering with the ability of the policy maker to fashion equilibrium outcomes.

In this paper, we analyze endogenous policy in a global coordination game. We consider a simple model of regime change that embeds many applications examined in the literature, including the currency crises game of Morris and Shin (1998). There is a large number of private agents (market participants), who may either attack the status quo (undertake actions that favor a regime change) or abstain from attacking. The status quo is abandoned if and only if the mass of agents attacking is sufficiently large. A policy maker, who is interested in influencing the probability of regime change, controls a policy instrument that affects the agents' payoff from attacking. Let θ parametrize the policy maker's willingness and ability to defend the status quo (the fundamentals). The policy maker moves first, setting the policy on the basis of her knowledge of θ, and agents move second, deciding whether to attack on the basis of their idiosyncratic noisy signals about θ and their observation of the policy choice. In the context of speculative currency attacks, the status quo is a

[2] See Morris and Shin (2001) for an extensive overview of the global-games literature. Earlier uniqueness results are also in Postlewaite and Vives (1987) and Chamley (1999).

currency peg, which the monetary authorities may attempt to defend by raising domestic interest rates or imposing capital controls; in industries with network externalities, a regime change can be interpreted as the adoption of a new technology standard, which the government may wish to encourage with appropriate subsidies or regulatory incentives; in the context of political change, the status quo corresponds to a dictator or some other sociopolitical establishment facing the risk of collapse if attacked by a large fraction of the society.

The main result of the paper (Theorem 1) establishes that the endogeneity of the policy leads to multiple equilibrium policies and multiple coordination outcomes, even when the fundamentals are observed with idiosyncratic noise. Different equilibria are sustained by different modes of coordination in the agents' response to the same policy choice, and such coordination is possible only because a policy intervention signals that the fundamentals are neither too weak nor too strong. If private agents coordinate on the same response to any level of the policy, the policy maker never finds it optimal to intervene. Hence, there always exists an *inactive policy equilibrium*, in which the policy maker sets the same policy for all θ and private agents play the Morris-Shin equilibrium in the continuation game. If instead private agents coordinate on a lenient continuation equilibrium (not attacking the status quo) when the policy maker raises the policy sufficiently high, and otherwise play an aggressive continuation equilibrium (attacking for sufficiently low private signals), the policy maker finds it optimal to intervene for an intermediate range of fundamentals. Hence, there also exists a *continuum of active-policy equilibria*, in which the policy maker raises the policy only for $\theta \in [\theta^*, \theta^{**}]$, and the regime collapses if and only if $\theta < \theta^*$. Which equilibrium is played, the level of the policy above which the regime is spared from an attack, and the threshold θ^* below which a regime change occurs, are all determined by self-fulfilling market expectations. This result thus manifests a kind of *policy traps*: In her attempt to fashion the equilibrium outcome, the policy maker reveals information that market participants can use to coordinate on multiple courses of action, and finds herself trapped in a position where both the optimal policy and the regime outcome are dictated by arbitrary market sentiments.

The second result of the paper (Theorem 2) establishes that information heterogeneity significantly reduces the equilibrium set as compared to common knowledge and enables meaningful predictions despite the existence of multiple equilibria. In particular, the probability of a regime change is monotonic in the fundamentals, the policy maker is "anxious to prove herself" with a costly policy intervention only for an intermediate region of fundamentals, and the "anxiety region" shrinks as the information in the market becomes precise. In essence, there is a *unique class* of equilibria. What is more, none of these predictions could be made if the fundamentals were common knowledge. Our results thus strike a delicate balance between equilibrium indeterminacy and predictive ability.

In the benchmark model, the policy maker faces no uncertainty about the aggressiveness of market expectations in any given equilibrium. We relax this assumption in Section 5 by introducing sunspots on which market participants may condition their response to the policy. The same level of the policy may now lead to different coordination outcomes along the same equilibrium. Applied to currency crises, these results help reconcile the fact documented by Kraay (2003) and others that raising interest rates or otherwise increasing the cost of attacking the currency does not systematically affect the likelihood or severity of a speculative attack – a fact that prima facie contradicts the policy prediction of Morris and Shin (1998). Moreover, these results help make sense of the Financial Times quote: The market may be equally likely to "interpret" a costly policy intervention either as a signal of strength, in which case the most desirable outcome may be attained, or as a signal of panic, in which case the policy maker's attempt to coordinate the market on the preferred course of action proves to be in vain.

On a more theoretical ground, this paper introduces *signaling* in *global coordination games*. The receivers (private agents) use the signal (policy) as a coordination device to switch between lenient and aggressive behavior in the global coordination game, thus inducing multiple equilibria in the signaling game. Our policy traps are thus different from the multiplicity of equilibria in standard signaling games, which is sustained by different systems of out-of-equilibrium beliefs and is not robust to proper forward induction refinements. They are also different from the multiplicity in standard global coordination games with exogenous public signals (Morris and Shin, 2001; Hellwig, 2002). The informational content of the policy is endogenous and is itself the result of the self-fulfilling expectations of the market. Most importantly, our result is about the multiplicity of equilibria in the policy (signaling) game, not the multiplicity in the coordination game. Finally, our policy traps are different from the *expectation traps* that originate in the government's lack of commitment (Obstfeld, 1996; Chari, Christiano and Eichenbaum, 1998; Albanesi, Chari and Christiano, 2002). Indeed, the policy maker need *not* have any incentive ex ante to commit to a particular policy, even if that would ensure a unique equilibrium.

The rest of the paper is organized as follows: Section 2 introduces the model and the equilibrium concept. Section 3 analyzes equilibria and Section 4 examines robust predictions. Section 5 introduces sunspots. Section 6 concludes. All formal proofs are confined to the Appendix.

2 The Model

2.1 Model Description

Consider an economy populated by a continuum of measure-one of private agents (also referred to as market participants), indexed by i and uniformly distributed over the $[0,1]$ interval. There are two possible regimes, a status quo and an alternative. Each agent can choose between two actions, either attack the status quo (take an action that favors a regime change) or abstain from attacking (take an action that favors the status quo). In addition, there is a policy maker who is interested in defending the status quo and controls a policy instrument that affects the agents' payoff from attacking.[3] We let $\theta \in \Theta$ parametrize the strength of the status quo, or equivalently the policy maker's willingness and ability to defend the status quo against a potential attack. θ is private information to the policy maker and corresponds to what Morris and Shin (1998) refer to as the "fundamentals". For simplicity, we let $\Theta = \mathbb{R}$ and the common prior shared by the private agents be a degenerate uniform.[4]

The game has three stages. In stage 1, the policy maker learns θ and sets the policy contingent on θ. In stage 2, each agent decides whether to attack after observing the policy choice and after receiving a noisy private signal $x_i = \theta + \varepsilon\xi_i$ about θ. The scalar $\varepsilon \in (0,\infty)$ parametrizes the precision of the agents' private information about θ and ξ_i is noise, i.i.d. across private agents and independent of θ, with absolutely continuous c.d.f. Ψ and density ψ strictly positive and continuously differentiable over the entire real line (unbounded full support) or a closed interval $[-1,+1]$ (bounded support). Finally, in stage 3, the policy maker observes the aggregate size of the attack, denoted by A, and decides whether to maintain or abandon the status quo.

The payoffs for a private agent who decides not to attack are normalized to zero. The payoff for an agent who decides to attack is $b > 0$ in the event the status quo is abandoned and $-c < 0$ otherwise. Letting a^i denote the probability agent i attacks the regime and D the probability the regime changes, agent i's payoff is given by $u_i = a_i[Db+(1-D)(-c)]$. Up to a linear transformation, this is equivalent to

$$u_i(a_i, D) = a^i(D - r),$$

where $r \equiv c/(b+c) \in (0,1)$ parametrizes the agent's cost of attacking. For simplicity, we assume

[3] We adopt the convention of female pronouns for the policy maker and masculine pronouns for the agents.

[4] Our results hold for any interval $\Theta \subseteq \mathbb{R}$ and any strictly positive and continuous density over Θ, provided that the game remains "global." We refer to Morris and Shin (2001) for a discussion of the role of degenerate uniform and general common priors in global coordination games.

the policy maker controls r directly. We let $\mathcal{R} = [\underline{r}, \overline{r}] \subset (0,1)$ be the domain of the policy, $C(r)$ the cost associated with raising the policy at level r, and $V(\theta, A)$ the net value of defending the status quo against an attack of size A. The policy maker's payoff is thus given by

$$U(r, A, D, \theta) = (1 - D)V(\theta, A) - C(r).$$

V is increasing in θ and decreasing in A, C is increasing in r, and both V and C are continuous. To simplify the exposition, we normalize $C(\underline{r}) = 0$ and let $V(\theta, A) = \theta - A$. Hence, the status quo is abandoned in stage 3 if and only if A exceeds θ. Define $\underline{\theta} \equiv 0$, $\overline{\theta} \equiv 1$, $\underline{x} \equiv \inf\{x : \Pr(\theta < \underline{\theta}|x) < 1\}$, and $\overline{x} \equiv \sup\{x : \Pr(\theta < \overline{\theta}|x) > 0\}$.[5] For $\theta < \underline{\theta}$ the collapse of the regime is inevitable and thus for $x < \underline{x}$ it is dominant for an agent to attack. Similarly, for $\theta > \overline{\theta}$ no attack can ever trigger a regime change and thus for $x > \overline{x}$ it is dominant not to attack. The interval $[\underline{\theta}, \overline{\theta}]$ thus represents the "critical range" of fundamentals for which the regime is sound but vulnerable to a sufficiently large attack.

2.2 Model Discussion and Applications

Many of the simplifying assumptions in the model are not essential for the results. For example, the policy maker may favor a regime change instead of the status quo. Similarly, the cost of the policy could depend on the fundamentals and the reaction of the market, and the value of preserving the status quo could depend on the level of the policy. Moreover, the policy maker must have private information about θ, but this information need not perfect. Finally, it is important that the continuation game following the policy choice is a coordination game, but the source of strategic complementarities is irrelevant. Indeed, although described as a three-stage game, the model essentially reduces to a two-stage game, where in the first stage the policy maker signals information about θ and in the second stage private agents play a global coordination game in response to the action of the policy maker.[6] In short, we suggest that our analysis may well fit a broad class of coordination environments. Below we discuss two concrete examples.

Currency Crises. Let the status quo be a currency peg, the private agents be speculators, and the policy maker be a central bank. Each speculator is endowed with a unit of wealth. Let

[5] Note that $\underline{x} = \underline{\theta} - \varepsilon$ and $\overline{x} = \overline{\theta} + \varepsilon$ if ξ has bounded support, whereas $\underline{x} = -\infty$ and $\overline{x} = +\infty$ if ξ has full support; if θ were common knowledge, $\underline{x} = \underline{\theta}$ and $\overline{x} = \overline{\theta}$.

[6] Stage 3 in our game serves only to introduce strategic complementarities in the actions of the private agents. All the results would hold true if the status quo were exogenously abandoned whenever A exceeds θ.

a_i be the fraction of wealth that speculator i converts to foreign currency and invests abroad (equivalently, the probability he attacks the peg) and A the size of the speculative attack. Denote with D the probability of devaluation, δ the domestic interest rate, δ^* the foreign interest rate, π the devaluation premium, and t the transaction cost associated with investing abroad. The expected payoff of a speculator is then given by

$$u_i = (1 - a_i)(1 + \delta) + a_i[D(1 + \delta^* + \pi - t) + (1 - D)(1 + \delta^* - t)].$$

Up to a linear transformation, this is equivalent to $u_i = a_i(D - r)$, where $r = (\delta - \delta^* + t)/\pi$. The monetary authority may increase r by raising domestic interest rates or transaction costs. Let $C(r)$ represent the cost of such policies, θ the bank's willingness and ability to defend the currency, and $V(\theta, A) = \theta - A$ the net value of maintaining the peg against an attack of size A.[7] The payoff of the bank is thus $U = (1 - D)V(\theta, A) - C(r)$. We conclude that speculative currency attacks map directly into the model of regime change of this paper. If r were exogenously fixed, the game would reduce to that of Morris and Shin (1998); if in addition θ were common knowledge, the game would reduce to that of Obstfeld (1986, 1996).

Revolutions and Political Reforms. There is a dictator (or some other sociopolitical establishment) who can be overturned only if a sufficiently large fraction of the population raises against the regime. If the revolution succeeds, the citizens who supported it enjoy the benefits associated with democracy and freedom, as well as the moral satisfaction of having done the right thing. Those who instead supported the dictator may free ride on some benefits but may also loose some rights or face sanctions from the new regime; let $b > 0$ denote the benefits, $m > 0$ the moral satisfaction, and $s > 0$ the potential sanctions. On the other hand, if the revolution is repressed, the dictator may imprison, torture, and punish the rebels, and may favor and reward the supporters; let p denote the penalties and q the rewards. The payoff of citizen i is thus

$$u_i = D\left[a_i(b + m) + (1 - a_i)(b - s)\right] + (1 - D)\left[a_i(-p) + (1 - a_i)q\right].$$

This is again strategically equivalent to $u_i = a_i(D - r)$, where $r = (p + q)/(p + q + s + m)$. Let $C(r)$ represent the cost of more sever punishments to the rebels or higher rewards to the supporters and $V(\theta, A) = \theta - A$ the net value of repressing a revolution of size A. The payoff of the dictator is thus $U = (1 - D)V(\theta, A) - C(r)$. We conclude that this game also maps into our model.

[7] For a more detailed discussion about the value of maintaining the peg and the costs of defensive policies, see Drazen (2001).

2.3 Equilibrium Definition

In the analysis that follows, we restrict attention to *perfect Bayesian equilibria* that are not sensitive to whether the idiosyncratic noise in the observation of the fundamentals has bounded or unbounded (full) support. We refer to equilibria that can be sustained under both specifications of the information structure as *robust equilibria.* This refinement imposes no restriction on the precision of private information and, as we explain in Section 4, only eliminates strategic effects that depend critically on the support of the signals.

Definition 1 *A* ***perfect Bayesian equilibrium*** *is a set of functions* $r, a, A, D, \mu,$ *such that:*

$$r(\theta) \in \arg\max_{r\in\mathcal{R}} U\left(r, A(\theta,r), D(\theta, A(\theta,r), r), \theta\right); \tag{1}$$

$$a(x,r) \in \arg\max_{a\in[0,1]} \int_{\Theta} u(r,a,D\left(\theta, A(\theta,r),r\right))d\mu\left(\theta|x,r\right) \text{ and } A(\theta,r) = \int_{-\infty}^{+\infty} a(x,r)\psi\left(\tfrac{x-\theta}{\varepsilon}\right)dx; \tag{2}$$

$$D(\theta, A, r) \in \arg\max_{D\in[0,1]} U(r,A,D,\theta) \quad \text{and} \quad D(\theta) = D(\theta, r(\theta), A(\theta, r(\theta))); \tag{3}$$

$$\mu\left(\theta|x,r\right) = 0 \text{ for all } \theta \notin \Theta(x) \text{ and } \mu\left(\theta|x,r\right) \text{ satisfies Bayes' rule for any } r \in r(\Theta(x)), \tag{4}$$

where $\Theta(x) \equiv \left\{\theta : \psi\left(\frac{x-\theta}{\varepsilon}\right) > 0\right\}$ *is the set of fundamentals* θ *consistent with signal* x.

Definition 2 *A perfect Bayesian equilibrium is* ***robust*** *if and only if the same policy* $r(\theta)$ *and the same probability of regime change* $D(\theta)$ *can be sustained with both bounded and unbounded idiosyncratic noise in the agents' observation of* θ.

$r(\theta)$ is the equilibrium policy, $\mu\left(\theta|x,r\right)$ the agent's posterior beliefs, $a(x,r)$ the agent's best response, $A(\theta,r)$ the corresponding size of an attack,[8] and $D(\theta)$ the equilibrium probability of regime change. Conditions (1), (2) and (3) guarantee that strategies are sequentially optimal, whereas condition (4) requires that beliefs never assign positive measure to fundamentals θ that are not compatible with signal x and are pinned down by Bayes' rule along the equilibrium path. In what follows, we also verify that all robust equilibria satisfy the *intuitive criterion*, first introduced in Cho and Kreps (1987). A perfect Bayesian equilibrium satisfies the intuitive criterion test if and only if no θ would be better off by choosing $r \neq r(\theta)$ should private agents' reaction not be sustained by out-of-equilibrium beliefs that assign positive measure to types for whom r is dominated in equilibrium; that is, if and only if $U\left(\theta\right) \geq U\left(r, A(\theta,r), D(\theta, A(\theta,r), r), \theta\right)$ for all θ, r,

[8] That $A(\theta,r) = \int_{-\infty}^{+\infty} a(x,r)\psi\left(\frac{x-\theta}{\varepsilon}\right)dx$ represents the fraction of agents attacking the status quo follows directly from the Law of Large Numbers when there are countable infinitely many agents; with a continuum, see Judd (1985).

and $A(\theta, r)$ satisfying (2) with $\mu \in \mathcal{M}(r)$, where $U(\theta)$ is the equilibrium payoff, $\Theta(r)$ is the set of θ for whom r is dominated in equilibrium, and $\mathcal{M}(r)$ is the set of beliefs that assign zero measure to any $\theta \in \Theta(r)$ whenever $\Theta(r) \subset \Theta(x)$.[9]

3 Policy Traps

3.1 Exogenous versus Endogenous Policy

Suppose for a moment that the policy is exogenously fixed, say at $r = \underline{r}$. Our model then reduces to that of Morris and Shin (1998), in which case iterated deletion of strictly dominated strategies selects a unique equilibrium profile and a unique system of equilibrium beliefs.

Proposition 1 (Morris-Shin) *In the global coordination game with exogenous policy, there exists a unique perfect Bayesian equilibrium, in which an agent attacks if and only if $x < x_{MS}$ and the status quo is abandoned if and only if $\theta < \theta_{MS}$. The thresholds x_{MS} and θ_{MS} are decreasing in $\underline{r}$ and are defined by*

$$\theta_{MS} = 1 - \underline{r} = \Psi\left(\frac{x_{MS} - \theta_{MS}}{\varepsilon}\right). \tag{5}$$

An agent finds it optimal to attack if and only if $\underline{r} \leq \mu(D(\theta) = 1|x) = 1 - \Psi\left(\frac{x - \theta_{MS}}{\varepsilon}\right)$. Thus, x_{MS} solves $\underline{r} = 1 - \Psi\left(\frac{x_{MS} - \theta_{MS}}{\varepsilon}\right)$. The fraction of private agents attacking is then $A(\theta) = \Psi\left(\frac{x_{MS} - \theta}{\varepsilon}\right)$. It follows that the status quo is abandoned if and only if $\theta < \theta_{MS}$, where θ_{MS} solves $\theta_{MS} = A(\theta_{MS}) = \Psi\left(\frac{x_{MS} - \theta}{\varepsilon}\right)$. Combining the above conditions gives (5).

The higher the cost $\underline{r}$, the less attractive it is for an agent to attack the status quo. It follows that x_{MS} and θ_{MS} are decreasing functions of $\underline{r}$, which suggests that the policy maker should be able to reduce the likelihood and severity of an attack simply by undertaking policies that reduce the individual payoff from attacking the regime. Indeed, that would be the end of the story if the policy did not convey any information to the market. However, since raising r is costly, the policy maker will not intervene if the collapse of the regime is inevitable, and hence any policy intervention signals that θ is not too low. On the other hand, as long as private agents do not attack when their private signals are sufficiently high, the policy maker faces at most a small attack when θ is sufficiently high, in which case there is no need to take costly policy measures. Therefore, any attempt to defend the status quo by raising r is correctly interpreted by the market as a signal of

[9]Formally, $\Theta(r) \equiv \{\theta$ such that $U(\theta) > U(r, A(\theta, r), D(\theta, A(\theta, r), r), \theta)$ for any $A(\theta, r)$ satisfying (2) with $\mu \in \mathcal{M}(r)\}$, where $\mathcal{M}(r) \equiv \{\mu$ satisfying (4) and such that $\mu(\theta|x, r) = 0$ for any $\theta \in \Theta(r)$ if $\Theta(r) \subset \Theta(x)\}$.

intermediate fundamentals,[10] in which case coordination on either an aggressive or a lenient course of action becomes possible. Different modes of coordination then create different incentives for the policy maker and result in different equilibrium policies.

Theorem 1 (Policy Traps) *In the global coordination game with endogenous policy, there exist multiple perfect Bayesian equilibria for any $\varepsilon > 0$.*

(a) There is an ***inactive-policy equilibrium:*** *The policy maker sets the policy at its cost-minimizing level $\underline{r}$ for all θ and the status quo is abandoned if and only if $\theta < \theta_{MS}$.*

(b) There is a continuum of ***active-policy equilibria:*** *Let $\tilde{r}$ solve $C(\tilde{r}) = 1 - \underline{r}$; for any $r^* \in (\underline{r}, \tilde{r}]$, there is an equilibrium in which the policy maker sets either $\underline{r}$ or r^*, the policy is raised at r^* only for $\theta \in [\theta^*, \theta^{**}]$, and the status quo is abandoned if and only if $\theta < \theta^*$, where*

$$\theta^* = C(r^*) \quad \text{and} \quad \theta^{**} = \theta^* + \varepsilon \left[\Psi^{-1}\left(1 - \tfrac{\underline{r}}{1-\underline{r}}\theta^*\right) - \Psi^{-1}\left(\theta^*\right) \right]. \tag{6}$$

The threshold θ^* is independent of ε and can take any value in $(\underline{\theta}, \theta_{MS}]$, whereas the threshold θ^{**} is increasing in ε and converges to θ^* as $\varepsilon \to 0$.

All the above equilibria are robust, satisfy the intuitive criterion, and can be supported by strategies for the agents that are monotonic in x.

The policy maker is thus subject to *policy traps*, which contrasts with the policy conjecture of Morris and Shin. In her attempt to use the policy so as to fashion the equilibrium outcome, the policy maker reveals information that facilitates coordination on multiple courses of action in the market and hence finds herself trapped in a situation where the optimal policy and the eventual fate of the regime are dictated by the arbitrary aggressiveness of the market expectations. The proof of Theorem 1 follows from Propositions 2 and 3, which we present in the next two sections.

3.2 Inactive Policy Equilibrium

With endogenous policy, the Morris-Shin outcome survives as a perfect-pooling equilibrium.

Proposition 2 (Perfect Pooling) *There is a robust inactive policy equilibrium, in which the policy maker sets $\underline{r}$ for all θ, private agents attack if and only if $x < x_{MS}$, independently of the level of the policy, and the status quo is abandoned if and only if $\theta < \theta_{MS}$. The thresholds x_{MS} and θ_{MS} are defined as in (5).*

[10] This interpretation of the information conveyed by an active policy intervention $r > \underline{r}$ follows from Bayes' rule if the observed level of the policy is on the equilibrium path, and from forward induction (the intuitive criterion) otherwise.

The construction of this equilibrium is illustrated in Figure 1. θ is on the horizontal axis, A and C on the vertical one. Since in equilibrium the policy maker sets $\underline{r}$ for all θ, the observation of $\underline{r}$ conveys no information about the fundamentals, in which case the continuation game starting after $r = \underline{r}$ is isomorphic to the Morris-Shin game. The threshold θ_{MS} is the point of intersection between the value of defending the status quo θ and the (cost of the) size of the attack $A(\theta, \underline{r})$. Let $\widetilde{r}$ solve $C(\widetilde{r}) = \theta_{MS}$. The equilibrium payoff is $U(\theta) = 0$ for all $\theta \leq \theta_{MS}$ and $U(\theta) = \theta - A(\theta, \underline{r}) > 0$ for all $\theta > \theta_{MS}$. Hence, any deviation to $r' > \widetilde{r}$ is dominated in equilibrium for all θ. Consider a deviation to some $r' \in (\underline{r}, \widetilde{r}]$ and let θ' and θ'' solve $\theta' = C(r') = A(\theta'', \underline{r})$. Note that $C(r') > \theta$ if and only $\theta < \theta'$ and $C(r') > A(\theta, \underline{r})$ if and only if $\theta > \theta''$, which implies that r' is dominated in equilibrium by $\underline{r}$ if and only if $\theta \notin [\theta', \theta'']$. Hence, if $\theta \in [\theta', \theta'']$ and the policy maker deviates to r', the market "learns" that $\theta \in [\theta', \theta'']$.[11] Since $\theta_{MS} \in [\theta', \theta'']$, one can construct beliefs that are compatible with the intuitive criterion and for which private agents continue to attack if and only if $x < x_{MS}$ for any $r > \underline{r}$, in which case it is pointless for the policy maker to ever raise the policy. (See Appendix for the specification of beliefs and the proof of robustness.)

Insert Figure 1 here

Clearly, any system of beliefs and strategies such that $A(\theta, r) \geq A(\theta, \underline{r})$ for all $r > \underline{r}$ sustains policy inaction as an equilibrium: If the size of the attack is non-decreasing in r, the policy maker can do no better than setting $r(\theta) = \underline{r}$ for all θ. Note also that a higher level of the policy increases the agents' cost of attacking the status quo and, other things equal, reduces the agents' incentive to attack. In an inactive-policy equilibrium, however, this payoff effect is offset by the higher probability agents attach to a regime change. The particular beliefs we consider in the proof of Proposition 2 have the property that these two effects just offset each other, in which case private agents use the same strategy on and off the equilibrium path and condition their behavior *only* on their private information. It is then as if the market does not pay any attention to the actions of the policy maker.[12]

[11] Throughout, the expression "the market learns X by observing Y" means that the event X becomes *common beliefs* among the agents given that Y is common knowledge; see Monderer and Samet (1989) and Kajii and Morris (1995).

[12] Moreover, private agents do not need to learn how to behave off the equilibrium path; they simply continue to play the same strategy they learned to play in equilibrium.

3.3 Active Policy Equilibria

The following proves the existence of robust active-policy equilibria in which the policy is raised at r^* for *every* $\theta \in [\theta^*, \theta^{**}]$.

Proposition 3 (Two-Threshold Equilibria) *For any $r^* \in (\underline{r}, \tilde{r}]$, there is a robust two-threshold equilibrium in which the policy maker sets r^* for all $\theta \in [\theta^*, \theta^{**}]$ and $\underline{r}$ otherwise; private agents attack if and only if either $r < r^*$ and $x < x^*$, or $x < \underline{x}$; finally, the status quo is abandoned if and only if $\theta < \theta^*$. The thresholds θ^* and θ^{**} are given by (6), and x^* solves $\mu(\theta < \theta^*|x^*, \underline{r}) = \underline{r}$. A two-threshold equilibrium exists if and only if $r^* \in (\underline{r}, \tilde{r}]$ or, equivalently, if and only if $\theta^* \in (\underline{\theta}, \theta_{MS}]$.*

The construction of a two-threshold equilibrium is illustrated in Figure 2. Take any $r^* \in (\underline{r}, \tilde{r}]$. Like in the inactive policy equilibrium, the observation of any policy choice $r > \underline{r}$ is interpreted by market participants as a signal of intermediate fundamentals. But contrary to the inactive-policy equilibrium, private agents switch from playing aggressively (attacking if and only if $x < x^*$) to playing leniently (attacking if and only if $x < \underline{x}$) whenever the policy maker sets $r \geq r^*$. Anticipating this reaction by market participants, the policy maker finds it optimal to raise the policy from $\underline{r}$ to r^* if and only if the fundamentals are strong enough that the net value of maintaining the status quo offsets the cost of the policy, i.e., $\theta \geq C(r^*)$, but not so strong that the cost of facing a small and unsuccessful attack is lower than the cost of a policy intervention, i.e., $A(\theta, \underline{r}) \geq C(r^*)$. The thresholds θ^* and θ^{**} are thus determined by the indifference conditions $C(r^*) = \theta^*$ and $C(r^*) = A(\theta^{**}, \underline{r}) = \Psi\left(\frac{x^* - \theta^{**}}{\varepsilon}\right)$. On the other hand, the threshold x^* solves the indifference condition of the agent, namely $\underline{r} = \mu(D(\theta) = 1|x^*, \underline{r})$, where

$$\mu(D(\theta) = 1|x^*, \underline{r}) = \mu(\theta < \theta^*|x^*, \underline{r}) = \frac{1 - \Psi\left(\frac{x^* - \theta^*}{\varepsilon}\right)}{1 - \Psi\left(\frac{x^* - \theta^*}{\varepsilon}\right) + \Psi\left(\frac{x^* - \theta^{**}}{\varepsilon}\right)}.$$

Combining the three indifference conditions gives (6). One can then construct out-of-equilibrium beliefs that satisfy the intuitive criterion and for which the strategy of the agents is sequentially optimal. (See Appendix for the specification of beliefs and the proof of robustness.)

`Insert Figure 2 here`

In any two-threshold equilibrium, the exact fundamentals never become common knowledge. What the market "learns" from equilibrium policy observations is only whether $\theta \in [\theta^*, \theta^{**}]$ or $\theta \notin [\theta^*, \theta^{**}]$, but this is enough to facilitate coordination on different courses of action.

For any r^*, the threshold θ^* below which the status quo is abandoned is independent of ε, whereas θ^{**} is increasing in ε and $\theta^{**} \to \theta^*$ as $\varepsilon \to 0$. As private information becomes more precise, the policy maker needs to intervene only for a smaller range of fundamentals. At the limit, the equlibrium policy has a spike at an arbitrary threshold $\theta^* \in (\underline{\theta}, \theta_{MS}]$ dictated by market expectations. Note also that a two-threshold equilibrium exists if and only if $\theta^* \leq \theta_{MS}$. The intuition behind this result is as follows. From the policy maker's indifference conditions, $\theta^* = C(r^*) = A(\theta^{**}, \underline{r})$, we infer that a higher r^* raises both the threshold θ^* and the size of the attack at θ^{**}. The latter is given by $\Pr(x \leq x^*|\theta^{**})$, which is also equal to $\Pr(\theta \geq \theta^{**}|x^*)$. Therefore, $\Pr(\theta \geq \theta^{**}|x^*)$ increases with r^*. For a marginal agent to be indifferent between attacking and non attacking conditional on $r = \underline{r}$, it must be that $\Pr(\theta \leq \theta^*|x^*)$ also increases with r^*. It follows that $\Delta\theta(r^*) \equiv \theta^{**} - \theta^*$ is decreasing in r^*. Obviously, $\Delta\theta(r^*)$ is also continuous in r^*. A two-threshold equilibrium exists if and only if $\Delta\theta(r^*) \geq 0$. By the monotonicity of $\Delta\theta(r^*)$, there exists at most one $\widetilde{r}$ such that $\Delta\theta(\widetilde{r}) = 0$, and $\Delta\theta(r^*) > 0$ if and only if $r^* < \widetilde{r}$. But $\Delta\theta(\widetilde{r}) = 0$ if and only if $\theta^* = \theta^{**}$, in which case the continuation game following $\widetilde{r}$ is (essentially) the Morris-Shin game and therefore $\theta^* = \theta^{**} = \theta_{MS}$. It follows that $\widetilde{r}$ solves $C(\widetilde{r}) = \theta_{MS} = 1 - \underline{r}$ and a two-threshold equilibrium exists if and only if $r^* \in (\underline{r}, \widetilde{r}]$, or equivalently, if and only if $\theta^* \in (\underline{\theta}, \theta_{MS}]$. In Section 4, we will establish that these properties extend to *all* robust equilibria of the game.

Finally, note that there are other strategies for the private agents that also sustain the same two-threshold equilibria. For example, we could have assumed private agents coordinate on the *most* aggressive continuation equilibrium (attack if and only if $x < \overline{x}$) whenever the policy does not meet market expectations ($r \neq \underline{r}, r^*$). Nonetheless, the beliefs and strategies we consider in Proposition 3 have the appealing property that the private agents follow the same aggressive behavior for any $r < r^*$ and the same lenient behavior for any $r \geq r^*$. It is then as if market participants simply "ignore" any attempt of the policy maker that falls short of market expectations and continue to play exactly as if there had been no intervention.

3.4 Discussion

We conclude this section with a few remarks about the role of coordination, signaling, and commitment in our environment.

First, consider environments where the market does *not* play a coordination game in response to the policy maker, such as when the policy maker (sender) interacts with a single representative agent (receiver).[13] In such environments, the policy can be non-monotonic if the receivers have access to exogenous information that allows to separate very high from very low types of the sender (Feltovich, Harbaugh and To, 2002). Moreover, multiple equilibria could possibly be supported by different out-of-equilibrium beliefs. However, a unique equilibrium would typically survive the intuitive criterion or other proper forward induction refinements. To the contrary, the multiplicity we have identified in this paper does not depend in any critical way on the specification of out-of-equilibrium beliefs, originates merely in endogenous coordination, and would not arise in environments with a single receiver. The comparison is sharp if we consider $\varepsilon \to 0$. The limit of all equilibria of the game with a single agent would have the latter attacking if and only if $x < \overline{\theta}$, and the status quo being abandoned if and only if $\theta < \overline{\theta}$. Contrast this with the result in Theorem 1, where the threshold θ^* below which the regime changes can take any value in $(\underline{\theta}, \theta_{MS}]$.

Second, consider environments where there is no signaling, such as standard global coordination games. As shown in Morris and Shin (2001) and Hellwig (2002), these games may exhibit multiple equilibria if market participants observe sufficiently informative *public* signals about the underlying fundamentals. The multiplicity of equilibria of Theorem 1, however, is substantially different from the kind of multiplicity in that literature. The policy in our model does generate a public signal about the fundamentals. Yet, the informational content of this signal is endogenous, as it depends on the particular equilibrium played in the market. Moreover, Theorem 1 is not about the possibility of multiple continuation equilibria in the coordination game that follows a given realization of the public signal; it is rather about how endogenous market coordination that is facilitated by the endogeneity of the policy leads to multiple equilibria in the policy game (the signaling game). In other words, it is the existence of *policy traps* the essence of Theorem 1.

Third, note that endogenous policy choices introduce a very specific kind of signal. The observation of $r > \underline{r}$ reveals that the fundamentals are neither too weak nor too strong, and it is this particular kind of information that restores the ability of the market to coordinate on different courses of action. In this respect, an interesting extension is to consider the possibility the policy itself is observed with noise. It can be shown that all equilibria of Theorem 1 are robust to the introduction of small bounded noise in the agents' observation of the policy, independently of whether the noise is aggregate or idiosyncratic.[14]

Lastly, consider the role of commitment. Note that policy traps arise in our environment *because* the policy maker moves first, revealing information which market participants use to coordinate

[13] See, for example, the seminal work by Spence (1973) or the currency-crises example by Drazen (2001).

[14] The proof of this claim is available upon request.

their response to the policy choice. This raises the question of whether the policy maker would be better off committing to a certain level of the policy before observing θ, thus inducing a unique continuation equilibrium in the coordination game. To see that commitment is not necessarily optimal, suppose the idiosyncratic noise ξ has full support and consider $\varepsilon \to 0$. If the policy maker commits ex ante to some level of the policy r, she incurs a cost $C(r)$ and ensures that the status quo will be abandoned if and only if $\theta < \widehat{\theta}(r) \equiv 1-r$; moreover, for all $\theta > \widehat{\theta}(r)$, $A(\theta,r) \to 0$ as $\varepsilon \to 0$.[15] Hence, the ex ante payoff from committing to r is $U(r) = \Pr(\theta \geq \widehat{\theta}(r))\mathbb{E}[\theta|\theta \geq \widehat{\theta}(r)] - C(r)$. Let $U_c = \max_r U(r)$ and $\theta_c = \min\{\widehat{\theta}(r_c)|r_c \in \arg\max_r U(r)\}$. If instead the policy maker retains the option to fashion the policy contingent on θ, the ex ante payoff depends on the particular equilibrium $(r^*, \theta^*, \theta^{**})$ the policy maker expects to be played.[16] As $\varepsilon \to 0$, $\theta^{**} \to \theta^*$ and $A(\theta, \underline{r}) \to 0$ for all $\theta \geq \theta^{**}$, meaning the policy maker pays $C(r^*)$ only for a negligible measure of θ. It follows that the ex ante value of discretion is $U_d = \Pr(\theta \geq \theta^*)\mathbb{E}\left[\theta|\theta \geq \theta^*\right]$. But note that $\theta_c > \underline{\theta}$ and therefore any $\theta^* \in (\underline{\theta}, \theta_c)$ necessarily leads to $U_d > U_c$. A similar argument holds for arbitrary ε. We conclude that, even when perfect commitment is possible, the policy maker will prefer discretion ex ante as long as she is not too pessimistic about future market sentiments.[17]

4 Robust Predictions

Propositions 2 and 3 left open the possibility that there also exist equilibria outside the two classes of Theorem 1, which would only strengthen our argument that policy endogeneity facilitates market coordination on multiple courses of action and leads to policy traps. Nonetheless, we are also interested in identifying equilibrium predictions that are not unduly sensitive to the particular assumptions about the underlying information structure of the game.

We first note that, when the noise has unbounded full support, for any $r^* \in (\underline{r}, \overline{r}]$ one can construct a one-threshold equilibrium, in which private agents threaten to attack the status quo whenever they observe any $r < r^*$, no matter how high their private signal x, thus forcing the

[15] These results follow from Proposition 1, replacing $\underline{r}$ with an arbitrary r.

[16] Note that the payoff for the policy maker associated with the pooling equilibrium of Proposition 2 equals the payoff associated with the two-threshold equilibrium in which $\theta^* = \theta_{MS}$.

[17] The above argument assumes commitment to a fixed uncontingent level of the policy. A similar argument, however, is true in the case of commitment to a contingent policy rule. If the policy maker is sufficiently optimistic about the aggressiveness of market expectations, she will continue to prefer discretion. Moreover, a contingent policy rule, unlike a fixed uncontingent policy, does not necessarily induce a unique continuation equilibrium in the coordination game. As a result, the optimal policy rule itself may be indeterminate, resulting in another sort of policy trap.

policy maker to raise the policy at r^* for all $\theta \geq \theta^*$, where $\theta^* = C(r^*)$. This threat is sustained by beliefs such that market participants interpret any failure of the policy maker to raise the policy as a signal that the status quo will be abandoned independently of their private information about the fundamentals. It seems more plausible, however, that market participants remain confident that the regime will survive when they observe sufficiently high x even if the policy maker fails to intervene, in which case all one-threshold equilibria disappear. Indeed, this is necessarily true when the noise is bounded: For $x > \overline{x} \equiv \overline{\theta} + \varepsilon$ private agents find it dominant not to attack, which implies that for all $\theta > \overline{x} + \varepsilon$ the policy maker faces no attack and hence sets $\underline{r}$.

On the other hand, when the noise has bounded support, it may be possible for the market to tell apart types that set the same policy. For example, suppose that the policy maker sets $r(\theta) = r'$ for any $\theta \in [\theta_1, \theta_2] \cup [\theta_3, \theta_4]$ and $r(\theta) \neq r'$ otherwise, where $\theta_1 < \theta_2 < \theta_2 + 2\varepsilon < \theta_3 < \theta_4$. Since $[\theta_1, \theta_2]$ and $[\theta_3, \theta_4]$ are sufficiently apart and the noise is bounded, whenever r' is observed in equilibrium, the market "learns" whether $\theta \in [\theta_1, \theta_2]$ or $\theta \in [\theta_3, \theta_4]$. In principle, the possibility for the market to separate different subsets of types who set the same policy may lead to equilibria different from the ones in Theorem 1. However, this possibility critically depends on small bounded noise and disappears with large bounded or unbounded supports, whatever the precision of the noise.

It is these considerations that motivated the refinement in Definition 2. Robustness only eliminates strategic effects that are unduly sensitive to whether the support of the signals is bounded or unbounded. The following then provides the converse to Theorem 1.

Theorem 2 (Robust Equilibria) *Every robust perfect Bayesian equilibrium belongs to one of the two classes of Theorem 1.*

The intuition for Theorem 2 is as follows. Consider any perfect Bayesian equilibrium of the game. If all θ set $\underline{r}$, we have perfect pooling. Otherwise, let

$$\theta' = \inf\{\theta : r(\theta) > \underline{r}\} \quad \text{and} \quad \theta'' = \sup\{\theta : r(\theta) > \underline{r}\}$$

be, respectively, the lowest and highest type who take a costly policy action. Since there is no aggregate uncertainty, the policy maker can perfectly anticipate whether the status quo will be abandoned when she sets r and hence is willing to pay the cost of a policy $r > \underline{r}$ only if this leads to no regime change. It follows that any equilibrium observation of $r > \underline{r}$ necessarily signals that the status quo will be maintained and induces any agent not to attack. If there were more than one r above $\underline{r}$ played in equilibrium, and the noise were unbounded, the policy maker could always save the status quo by setting the lowest equilibrium r above $\underline{r}$. Since $C(r)$ is strictly increasing, it follows that at most one r above $\underline{r}$ is played in any robust equilibrium. Let r^* denote this level of the policy and define θ^* and θ^{**} as in Theorem 1. Obviously, it never pays to raise the policy for any $\theta < \theta^*$. Hence, in any robust equilibrium, $\theta' \geq \theta^*$. If the noise were bounded, all $\theta > \overline{x} + \varepsilon = \overline{\theta} + 2\varepsilon$ would necessarily set $\underline{r}$. Hence, in any robust equilibrium, $\theta'' < \infty$. Compare now the strategy of the private agents in any such equilibrium with the strategy in the corresponding two-threshold equilibrium. Note that a regime change never occurs for $\theta > \theta^*$ as the policy maker can guarantee herself a positive payoff by setting $r = r^*$. If the status quo is preserved also for some $\theta < \theta^*$, then the incentives to attack when observing $\underline{r}$ are lower than when a regime change occurs for *all* $\theta < \theta^*$. Similarly, if the policy maker does not raise the policy at r^* for some $\theta \in [\theta^*, \theta'']$, then the observation of $r = \underline{r}$ is less informative of regime change than in the case where $r(\theta) = r^*$ for *all* $\theta \in [\theta^*, \theta'']$. Hence, private agents are most aggressive at $\underline{r}$ when $D(\theta) = 1$ for all $\theta < \theta^*$ and $r(\theta) = r^*$ for all $\theta \in [\theta^*, \theta'']$, which, by definition of θ^{**}, is possible if and only if $\theta'' = \theta^{**}$. Equivalently, the size of the attack $A(\theta, \underline{r})$ in the two-threshold equilibrium corresponding to r^* represents an upper bound on the size of the attack in *any* active-policy equilibrium in which r^* is played. It follows that, in any robust equilibrium, $\theta'' \leq \theta^{**}$. Since $\theta^{**} < \theta^*$ whenever $r^* > \widetilde{r}$, this immediately rules out the possibility of equilibria in which $r^* > \widetilde{r}$. On the other hand, for any $r^* \leq \widetilde{r}$, we have

$$\theta^* \leq \theta' \leq \theta'' \leq \theta^{**}.$$

It follows that the anxiety region of any robust equilibrium is bounded by the anxiety region of the corresponding two-threshold equilibrium. By iterated deletion of strictly dominated strategies, one can also show that the status quo is necessarily abandoned for all $\theta < \theta^*$ and that $\theta' = \theta^* < \theta_{MS}$, which completes the proof of Theorem 2. (See Appendix.)

Theorem 2 permits us to make robust predictions: First, the probability of regime change is monotonic in the fundamentals. Second, the probability of regime change is lower in any active-policy equilibrium than under inactive policy ($\theta^* \leq \theta_{MS}$). Third, when the fundamentals are either very weak or very strong, private agents can easily recognize this, in which case there is no value for the policy maker to intervene; it is then only for a small range of moderate fundamentals that the agents are likely to be "uncertain" or "confused" about the eventual fate of the regime, and it is thus only for moderate fundamentals that the policy maker is "anxious" to undertake a costly policy action. Fourth, the "anxiety region" shrinks as the precision of the agents' information increases. In the limit, the policy converges to a spike around the threshold below which a regime change occurs.[18]

Corollary *The limit of any robust equilibrium as $\varepsilon \to 0$ is such that the policy is $r(\theta) = \underline{r}$ for all $\theta \neq \theta^*$, for any arbitrary $\theta^* \in (\underline{\theta}, \theta_{MS}]$, and the status quo is abandoned if and only if $\theta < \theta^*$.*

The above predictions are intuitive. Nevertheless, it would have been impossible to make any of these predictions if the fundamentals were common knowledge.

Proposition 4 (Common Knowledge) *Suppose $\varepsilon = 0$. A policy $r(\theta)$ can be part of a subgame perfect equilibrium if and only if $C(r(\theta)) \leq \theta$ for $\theta \in [\underline{\theta}, \overline{\theta}]$ and $r(\theta) = \underline{r}$ for $\theta \notin [\underline{\theta}, \overline{\theta}]$. Similarly, a regime outcome $D(\theta)$ of any shape over $[\underline{\theta}, \overline{\theta}]$ can be part of a subgame perfect equilibrium.*

That is, if θ were common knowledge, the equilibrium policy $r(\theta)$ could take essentially any shape in the critical range $[\underline{\theta}, \overline{\theta}]$. For example, it could have multiple discontinuities and multiple non-monotonicities. The equilibrium probability of regime change $D(\theta)$ could also take any shape in $[\underline{\theta}, \overline{\theta}]$ and need not be monotonic. For example, the status quo could be abandoned for any subset of $[\underline{\theta}, \overline{\theta}]$. These results contrast sharply with our results in Theorem 2. We conclude that introducing idiosyncratic noise in the observation of the fundamentals *does* reduce significantly the equilibrium set as compared to the common knowledge case. The global-game methodology thus maintains a strong selection power even in our multiple-equilibria environment and enables meaningful predictions that would have been impossible otherwise.

Remark Theorem 2 leaves open the possibility that there exist active-policy equilibria different from the two-threshold equilibria of Proposition 3, namely equilibria in which the policy is active only for a subset of $[\theta^*, \theta^{**}]$. This possibility, however, does not affect any of the predictions mentioned above and can be ruled out by imposing monotonicity of the strategy of the agents in their private information. Monotonicity is guaranteed when the noise satisfies the monotone likelihood ratio property (MLRP), that is, when $\psi'(\xi)/\psi(\xi)$ is decreasing in ξ. (See Appendix.)

[18] An interesting implication of the Corollary is that, in the limit, all active-policy equilibria are observationally equivalent to the inactive-policy equilibrium in terms of the level of the policy, although they are very different in terms of the regime outcome. An econometrician may then fail to predict the probability of regime change on the basis of information on the fundamentals and the policy. An even sharper dependence of observable outcomes on unobservable market sentiments arises if one introduces sunspots, as we discuss in the next section.

5 Uncertainty over the Aggressiveness of Market Expectations

The analysis so far assumed the policy maker was able to anticipate perfectly the aggressiveness of market expectations, which we identify with the threshold r^* at which private agents switch from an aggressive to a lenient response to the policy. In reality, however, market expectations are hard to predict, even when the underlying economic fundamentals are perfectly known to the policy maker. To capture this kind of uncertainty, we introduce payoff-irrelevant sunspots, on which private agents may condition their reaction to the policy maker. Instead of modelling explicitly the sunspots, we assume directly that r^* is a random variable with c.d.f. Φ over a compact support $\mathcal{R}^* \subseteq \mathcal{R}$. The realization of r^* is common knowledge among the private agents, but is unknown to the policy maker when she sets the policy, in which case random variation in r^* leads to random variation in the probability of regime change for given policy. Different sunspot equilibria are associated with different distributions $(\Phi, \mathcal{R}^*)$ and result in different equilibrium policies $r(\theta)$.

Proposition 5 (Sunspot Equilibria) *Take any random variable r^* with compact support $\mathcal{R}^* \subseteq (\underline{r}, \widetilde{r})$ and distribution Φ. For $\varepsilon > 0$ sufficiently small,*[19] *there exist thresholds $\theta^* \in (\underline{\theta}, \theta_{MS})$ and $\theta^{**} \in (\theta^*, \overline{\theta})$ and a robust equilibrium such that: The policy maker sets $r(\theta) = \underline{r}$ for $\theta \notin [\theta^*, \theta^{**}]$, and $r(\theta) \in \mathcal{R}^*$ with $r(\theta)$ non-decreasing in θ for $\theta \in [\theta^*, \theta^{**}]$: the status quo is abandoned with certainty for $\theta < \theta^*$, with probability less than one and non-increasing in θ for $\theta \in [\theta^*, \theta^{**}]$, and is never abandoned for $\theta > \theta^{**}$. Finally, θ^* is independent of ε, whereas $\theta^{**} \to \theta^*$ as $\varepsilon \to 0$.*

The sunspot equilibria of Proposition 5 are qualitatively similar to the two-threshold equilibria of Proposition 3. The probability of regime change is monotonic in the fundamentals, the policy is active only for an intermediate range of fundamentals, and this range vanishes as $\varepsilon \to 0$. $\mathcal{R}^*$ represents the set of random thresholds r^* such that private agents coordinate on an aggressive response whenever $r < r^*$ and on a lenient one whenever $r \geq r^*$. When $\theta < \theta^*$, raising the policy to any level in $\mathcal{R}^*$ is too costly compared to the expected value of defending the status quo, in which case the policy maker finds it optimal to set $\underline{r}$ and a regime change occurs with certainty. When instead $\theta \in [\theta^*, \theta^{**}]$, it pays to raise the policy at some level in $\mathcal{R}^*$ so as to lower the probability of a regime change. Since the value from defending the status quo is increasing in θ, so is the optimal policy in the range $[\theta^*, \theta^{**}]$. Finally, for $\theta \geq \theta^{**}$, the size of the attack at $\underline{r}$ is so small that the policy maker prefers the cost of such an attack to the cost of a policy intervention. The

[19] The assumption that ε is sufficiently small is not essential; it ensures $\theta^{**} < \overline{\theta}$, which we use only to simplify the construction of these equilibria (see the Appendix).

thresholds θ^* and θ^{**} are again given by the relevant indifference conditions for the policy maker, but differ from the ones we derived in the absence of sunspots. The definition of the thresholds and the complete proof of the above proposition are provided in the Appendix.[20]

With random variation in the aggressiveness of market expectations, policy traps take an even stronger form. Not only the policy maker has to adopt a policy that confirms market expectations, but also the coordination outcome that follows a policy intervention depends on market sentiments. For example, commentaries in the media that provide no information about the fundamentals and only second-guess the reaction of the market, may interfere with the effectiveness of policy intervention.

In the context of currency crises, empirical evidence suggests that raising the level of domestic interest rates or taking other defense measures does not systematically prevent an exchange-rate collapse. Kraay (1993), for example, studies the behavior of interest rates and other measures of monetary policy during 192 episodes of successful and unsuccessful speculative currency attacks and finds a "striking lack of any systematic association whatsoever between interest rates and the outcome of speculative attacks," even after controlling for various observable fundamentals. This evidence is hard to reconcile with Morris and Shin's (1998) prediction that defense policies decrease the probability of devaluation, but is consistent with the predictions of Proposition 5, where the same combination of interest rates and fundamentals may lead in equilibrium to either no devaluation or a collapse of the currency.

Finally, the results of this section help understand and formalize the kind of arguments that commonly appear in the popular press, like the one we quoted from Financial Times in the beginning of the paper. Once the policy maker has intervened in an attempt to achieve a favorable outcome, the market may be equally likely to "interpret" this action either as a signal of strength, in which case market participants coordinate on the desirable course of action, or as a signal of panic, in which case the policy maker's attempt proves in vain. And faced with the need to guess the largely arbitrary reaction of the market, the policy maker finds it even harder to decide what the optimal policy is.

---

[20] If one takes a sequence of sunspot equilibria such that $\mathcal{R}^*$ converges to a single point r^*, the thresholds θ^* and θ^{**} in Proposition 5 converge to the ones given in (6). That is, the two-threshold equilibria of Proposition 3 can be read as the limit of sunspot equilibria.

6 Concluding Remarks

In this paper we investigated the ability of a policy maker to influence market expectations and control equilibrium outcomes in economies where agents play a global coordination game. We found that policy endogeneity leads to policy traps, where the policy maker is forced to conform to the largely arbitrary expectations of the market instead of being able to fashion the equilibrium outcome. There is an inactive-policy equilibrium in which market participants coordinate on "ignoring" any attempt of the policy maker to affect market behavior, as well as a continuum of active-policy equilibria in which market participants coordinate on the level of the policy beyond which they "reward" the policy maker by playing a favorable continuation equilibrium. Despite equilibrium multiplicity, information heterogeneity significantly reduces the equilibrium set as compared to the case of common knowledge and enables meaningful predictions.

Stabilization policy in economies with macroeconomic complementarities, preemptive policies against speculative currency attacks, and regulatory intervention during financial crises, are only a few examples where our results might be relevant. We leave these applications and the extension of our results to more complex environments for future research.

Appendix

Proof of Proposition 1. The characterization of the equilibrium in our setting is presented in the main text. Uniqueness is established in Morris and Shin (1998). ■

Proof of Theorem 1. Follows from Propositions 2 and 3. ■

Proof of Proposition 2. Since in equilibrium the policy maker sets $\underline{r}$ for all θ, the observation of $\underline{r}$ conveys no information about θ and hence the continuation game starting after the policy maker sets $r = \underline{r}$ is isomorphic to the Morris-Shin game. It follows that there is a unique continuation equilibrium, in which an agent attacks if and only if $x < x_{MS}$ and the status quo is abandoned if and only if $\theta < \theta_{MS}$. Equilibrium beliefs are then pinned down by Bayes' rule.

Next, consider out-of-equilibrium levels of the policy. Note that $\tilde{r}$ (defined in Theorem 1) solves $C(\tilde{r}) = \theta_{MS} = 1 - \underline{r}$. Any $r > \tilde{r}$ is dominated in equilibrium by $\underline{r}$ for all θ: For $\theta < \theta_{MS}$,

$V(\theta,0)-C(r)<0$; for $\theta\geq\theta_{MS}$, $C(r)>\theta_{MS}>A(\theta,\underline{r})$ and thus $V(\theta,0)-C(r)<V(\theta,A(\theta,\underline{r}))$. On the other hand, any $r\in(\underline{r},\widetilde{r})$ is dominated in equilibrium by $r(\theta)$ for any θ such that $\theta<C(r)$ or $A(\theta,\underline{r})<C(r)$, where $C(r)\leq C(\widetilde{r})=\theta_{MS}$. Therefore, for any out-of-equilibrium $r\neq\widetilde{r}$, one can construct out-of-equilibrium beliefs $\mu\in\mathcal{M}(r)$ that satisfy (4) and such that $\mu(\theta<\theta_{MS}|x,r)$ is non-increasing in x and satisfies $\mu(\theta<\theta_{MS}|x,r)=r$ at $x=x_{MS}$. For such beliefs, a agent finds it optimal to attack if and only if $x<x_{MS}$, in which case the status quo is abandoned if and only if $\theta<\theta_{MS}$. Finally, note that $\widetilde{r}$ is dominated in equilibrium by $\underline{r}$ for all $\theta\neq\theta_{MS}$ and let $\mu(\theta_{MS}|x,\widetilde{r})=1$ for all x such that $\theta_{MS}\in\Theta(x)$, and $\mu(\theta|x,\widetilde{r})=\mu(\theta|x)$ otherwise, so that again $\mu\in\mathcal{M}(r)$ and (4) is satisfied. It follows that there is a mixed-strategy equilibrium for the continuation game following $r=\widetilde{r}$, in which an agent attacks if and only if $x<x_{MS}$ and type θ_{MS} abandons the status quo with probability $\widetilde{r}$.

Given that private agents attack if and only if $x<x_{MS}$ for any r, it is optimal for the policy maker to set $r(\theta)=\underline{r}$ for all θ. Finally, consider robustness in the sense of Definition 2. Given $\theta_{MS}=1-\underline{r}\in(0,1)$, (5) is satisfied for any $\varepsilon>0$ and any c.d.f. Ψ, with either bounded or unbounded support, by simply letting $x_{MS}=\theta_{MS}+\varepsilon\Psi^{-1}(\theta_{MS})$. It follows that the policy $r(\theta)=\underline{r}$ for all θ, and the probability of regime change $D(\theta)=1$ for $\theta<\theta_{MS}$ and $D(\theta)=0$ otherwise, can be sustained as a robust equilibrium. ■

Proof of Proposition 3. Let $\widehat{\theta}\in[\theta^*,\theta^{**}]$ solve $V(\widehat{\theta},A(\widehat{\theta},\underline{r}))=0$, where $A(\widehat{\theta},\underline{r})=\Psi\left(\frac{x^*-\widehat{\theta}}{\varepsilon}\right)$, and note that, for any $r<r^*$, private agents trigger a regime change if and only if $\theta<\widehat{\theta}$.

Consider first the behavior of the agents. When $r=\underline{r}$, beliefs are pinned down by Bayes' rule for all x, since (6) ensures $\Theta(x)\not\subseteq[\theta^*,\theta^{**}]$ for all x.[21] Therefore,

$$\mu(D(\theta)=1|x,\underline{r})=\mu(\theta<\widehat{\theta}|x,\underline{r})=\mu(\theta<\theta^*|x,\underline{r})=\frac{1-\Psi\left(\frac{x-\theta^*}{\varepsilon}\right)}{1-\Psi\left(\frac{x-\theta^*}{\varepsilon}\right)+\Psi\left(\frac{x-\theta^{**}}{\varepsilon}\right)},$$

which is decreasing in x. Let x^* be the unique solution to $\mu(\theta<\theta^*|x^*,\underline{r})=\underline{r}$. For any $r\in(\underline{r},r^*)$, we consider out-of-equilibrium beliefs μ such that $\mu(\theta<\widehat{\theta}|x,r)$ is non-increasing in x and $\mu(\theta<\widehat{\theta}|x,r)=r$ at $x=x^*$. Furthermore, if for some $\theta\in\Theta(x)$, r is not dominated in equilibrium by $r(\theta)$, i.e., if $C(r)\leq\min\{A(\theta,\underline{r}),\theta\}$, we further restrict μ to satisfy $\mu(\theta|x,r)=0$ for all $\theta\in\Theta(x)$ such that $\theta<C(r)$ or $A(\theta,\underline{r})<C(r)$. When $r=r^*$, Bayes' rule implies $\mu(\theta\in[\theta^*,\theta^{**}]|x,r^*)=1$ for all x such that $\Theta(x)\cap[\theta^*,\theta^{**}]\neq\varnothing$. Finally, for any (x,r) such that either $r=r^*$ and

[21] When the noise is unbounded, this is immediate; when it is bounded, it follows from the fact that $|\theta^{**}-\theta^*|<2\varepsilon$ and $\Theta(x)=[x-\varepsilon,x+\varepsilon]$, for any x.

$\Theta(x) \cap [\theta^*, \theta^{**}] = \varnothing$, or $r > r^*$, we let $\mu(\theta \geq \underline{\theta}|x, r) = 1$ for all $x \geq \underline{x}$ and $\mu(\theta \geq \underline{\theta}|x, r) = 0$ otherwise. Given these beliefs, (4) is satisfied, $\mu \in \mathcal{M}(r)$ for any r, and the strategy of the agents is sequentially optimal.

Consider next the policy maker. Given the strategy of the agents, the policy maker clearly prefers $\underline{r}$ to any $r \in (\underline{r}, r^*)$ and r^* to any $r > r^*$. θ^* is indifferent between setting r^* (in which case the regime is not attacked) and setting $\underline{r}$ (in which case the status quo is abandoned) if and only if $V(\theta^*, 0) - C(r^*) = 0$, i.e. $\theta^* = C(r^*)$. Similarly, θ^{**} is indifferent between setting r^* (and not being attacked) and setting $\underline{r}$ (and being attacked without being forced to abandon the status quo) if and only if $V(\theta^{**}, 0) - C(r^*) = V(\theta^{**}, A(\theta^{**}, \underline{r}))$, i.e. $C(r^*) = A(\theta^{**}, \underline{r})$, where $A(\theta^{**}, \underline{r}) = \Psi\left(\frac{x^* - \theta^{**}}{\varepsilon}\right)$. For any $\theta < \theta^*$, $\underline{r}$ is optimal; for any $\theta \in (\theta^*, \theta^{**})$, $A(\theta, \underline{r}) > A(\theta^{**}, \underline{r}) = C(r^*)$ and thus r^* is preferred to $\underline{r}$, whereas the reverse is true for any $\theta > \theta^{**}$. From the indifference conditions for θ^*, θ^{**}, and x^*, we obtain $x^* = \theta^{**} + \varepsilon\Psi^{-1}(\theta^*)$ and

$$\theta^{**} = \theta^* + \varepsilon\left[\Psi^{-1}\left(1 - \tfrac{\underline{r}}{1-\underline{r}}\theta^*\right) - \Psi^{-1}(\theta^*)\right]. \tag{7}$$

It follows that $\theta^* \leq \theta^{**}$ if and only if $\theta^* \leq 1 - \underline{r} = \theta_{MS}$. Using $\theta^* = C(r^*)$ and $\theta_{MS} = C(\widetilde{r})$, we infer that a two threshold equilibrium exists if and only if $\theta^* \in (\underline{\theta}, \theta_{MS}]$, or equivalently $r^* \in (\underline{r}, \widetilde{r}]$. Finally, consider robustness in the sense of Definition 2. For any $r^* \in (\underline{r}, \widetilde{r})$, let $\theta^* = C(r^*)$ and take any arbitrary $\theta^{**} > \theta^*$. For any c.d.f. Ψ with either bounded support $[-1, +1]$ or unbounded full support $\mathbb{R}$, we have that $\kappa \equiv \left[\Psi^{-1}\left(1 - \frac{\underline{r}}{1-\underline{r}}\theta^*\right) - \Psi^{-1}(\theta^*)\right] \in (0, \infty)$ since $(1 - \frac{\underline{r}}{1-\underline{r}}\theta^*) \in (\theta^*, 1)$. Hence, for any $\theta^{**} > \theta^*$ and for any κ, condition (7) can always be satisfied by letting $\varepsilon = \frac{\theta^{**} - \theta^*}{\kappa} \in (0, \infty)$. For $r^* = \widetilde{r}$, $\theta^* = \theta^{**} = \theta_{MS}$ and (7) holds for every ε and every Ψ. The indifference condition for the agents is then clearly satisfied at $x^* = \theta^{**} + \varepsilon\Psi^{-1}(\theta^*)$. It follows that the policy $r(\theta) = r^*$ if $\theta \in [\theta^*, \theta^{**}]$ and $r(\theta) = \underline{r}$ otherwise, and the probability of regime change $D(\theta) = 1$ if $\theta < \theta^*$ and $D(\theta) = 0$ otherwise, can be sustained as a robust equilibrium. ■

Proof of Theorem 2. When no r other than $\underline{r}$ is played in equilibrium, we have the pooling equilibrium in (a) of Theorem 1. Hence, in what follows, we consider equilibria in which $r(\theta) > \underline{r}$ for some θ and we let

$$\theta' = \inf\{\theta : r(\theta) > \underline{r}\} \quad \text{and} \quad \theta'' = \sup\{\theta : r(\theta) > \underline{r}\}.$$

We prove the result in a sequence of three lemmas.

Lemma 1. In any robust equilibrium, there is at most one level of the policy $r^* > \underline{r}$ played in equilibrium and $\theta'' < \infty$.

Proof. Any $r > \underline{r}$ is played in equilibrium only if it leads to no regime change, i.e. only if $D(\theta) = 0$; all θ who abandon the status quo set $r = \underline{r}$. Agents attack if and only if the probability of regime change exceeds the cost of attacking, i.e. $\int_\Theta D(\theta)\, d\mu(\theta|x, r) \geq r$. If the noise is unbounded, any equilibrium $r > \underline{r}$ results in a posterior $\mu(\Theta_0|x, r) = 1$ for all x, where $\Theta_0 \equiv \{\theta : D(\theta) = 0\}$ is the set of fundamentals for which the status quo is maintained. Hence, no agent ever attacks when he observes an equilibrium $r > \underline{r}$, and thus $A(\theta, r) = 0$ for any equilibrium $r > \underline{r}$. Since $C(r)$ is strictly increasing in r, this also implies that, in any robust equilibrium, at most one level of the policy $r^* \neq \underline{r}$ will be chosen by the policy maker, which proves the first part of the lemma. Next, if the noise were bounded, it would be dominant for an agent not to attack whenever $x > \overline{x} \equiv \overline{\theta} + \varepsilon$, in which case $A(\theta, r) = 0$ for all r whenever $\theta > \overline{x} + \varepsilon$ and thus $r(\theta) = \underline{r}$ for all $\theta > \overline{x} + \varepsilon$. Therefore, an equilibrium with $\theta'' = \infty$ could never be sustained with bounded noise, which proves the second part of the lemma. □

For the rest of the proof, it suffices to consider unbounded noise. Given any $r^* \in (\underline{r}, \overline{r}]$, define the thresholds

$$\theta^* = C(r^*) \quad \text{and} \quad \theta^{**} = \theta^* + \varepsilon \left[\Psi^{-1}\left(1 - \tfrac{\underline{r}}{1-\underline{r}}\theta^*\right) - \Psi^{-1}(\theta^*) \right],$$

and note that $\theta^{**} > \theta^*$ when $\theta^* < 1 - \underline{r} = \theta_{MS}$, i.e. for $r^* < \widetilde{r} = C^{-1}(\theta_{MS})$, $\theta^{**} = \theta^*$ when $r^* = \widetilde{r}$, and $\theta^{**} < \theta^*$ when $r^* > \widetilde{r}$.

Lemma 2. For any $r^* \in (\underline{r}, \overline{r}]$, $\theta^* \leq \theta'$ and $\theta'' \leq \theta^{**}$.

Proof. Since for any $\theta < \theta^*$, r^* is strictly dominated by $\underline{r}$, it immediately follows that $\theta' \geq \theta^*$. On the other hand, any $\theta > \theta^*$ can always set r^*, face no attack, and ensure a payoff $V(\theta, 0) - C(r^*) > 0$. Therefore, necessarily $D(\theta) = 0$ for all $\theta > \theta^*$. However, there may exist $\theta < \theta^*$ that also do not abandon the status quo in equilibrium. Define $\delta(x)$ as the probability, conditional on x, that $\theta < \theta^*$ and $D = 0$. Further, define $p(x)$ as the probability, conditional on x, that $\theta \in [\theta^*, \theta'']$ and $r(\theta) = \underline{r}$. Then, the probability of regime change conditional on x and $\underline{r}$ is given by

$$\mu(D = 1|x, \underline{r}) = \frac{1 - \Psi\left(\frac{x-\theta^*}{\varepsilon}\right) - \delta(x)}{1 - \Psi\left(\frac{x-\theta^*}{\varepsilon}\right) + p(x) + \Psi\left(\frac{x-\theta''}{\varepsilon}\right)}.$$

Clearly, an agent does not attacks at $\underline{r}$ whenever $\mu(D = 1|x, \underline{r}) < \underline{r}$. Define

$$F(x; \theta^*, \theta'') \equiv \frac{1 - \Psi\left(\frac{x-\theta^*}{\varepsilon}\right)}{1 - \Psi\left(\frac{x-\theta^*}{\varepsilon}\right) + \Psi\left(\frac{x-\theta''}{\varepsilon}\right)} = \left[1 + \frac{\Psi\left(\frac{x-\theta''}{\varepsilon}\right)}{1 - \Psi\left(\frac{x-\theta^*}{\varepsilon}\right)}\right]^{-1}.$$

Note that $F(x;\theta^*,\theta'')$ is strictly decreasing in x, and let $\widehat{x}(\theta^*,\theta'')$ solve

$$F(\widehat{x};\theta^*,\theta'') = \underline{r}. \tag{8}$$

Since $\delta(x) \geq 0$ and $p(x) \geq 0$, we have $\mu(D=1|x,\underline{r}) \leq F(x;\theta^*,\theta'')$ for all x. It follows that, whenever $x > \widehat{x}(\theta^*,\theta'')$,

$$\mu(D=1|x,\underline{r}) \leq F(x;\theta^*,\theta'') < F(\widehat{x};\theta^*,\theta'') = \underline{r}$$

and therefore any agent with $x > \widehat{x}(\theta^*,\theta'')$ does not attack in equilibrium. That is, in any equilibrium in which r^* is played, private agents are necessarily at most as aggressive as they would be if it were the case that $D(\theta) = 1$ for all $\theta < \theta^*$ and $r(\theta) = r^*$ for all $\theta \in [\theta^*,\theta'']$. This is intuitive for (i) a positive probability that $D(\theta) = 0$ for some $\theta < \theta^*$ reduces the incentives to attack for every x, and (ii) a positive probability that $r(\theta) = \underline{r}$ for some $\theta \in [\theta^*,\theta'']$ makes the observation of $\underline{r}$ less informative of regime change and therefore also reduces the incentives to attack conditional on $\underline{r}$ and x. It follows that, for any θ,

$$A(\theta,\underline{r}) \leq \Psi\left(\frac{\widehat{x}(\theta^*,\theta'') - \theta}{\varepsilon}\right).$$

From the indifference condition for θ'', we have that $A(\theta'',\underline{r}) = C(r^*)$. Since $C(r^*) = \theta^*$, it follows that $\theta^* \leq \Psi\left(\frac{\widehat{x}(\theta^*,\theta'')-\theta''}{\varepsilon}\right)$. Using $\widehat{x}(\theta^*,\theta^{**}) = x^*$ and the indifference condition for the two-threshold equilibrium of Proposition 3 $\Psi\left(\frac{x^*-\theta^{**}}{\varepsilon}\right) = C(r^*) = \theta^*$ we have that in any equilibrium in which r^* is played, θ'' must satisfy

$$\widehat{x}(\theta^*,\theta^{**}) - \theta^{**} \leq \widehat{x}(\theta^*,\theta'') - \theta''.$$

From (8), $\widehat{x}(\theta^*,\theta'') - \theta''$ is decreasing in θ'', and hence we conclude that $\theta'' \leq \theta^{**}$. □

Recall that $\theta^* \leq \theta^{**}$ if and only if $r^* \leq \widetilde{r} \equiv C^{-1}(\theta_{MS})$. For any $r^* > \widetilde{r}$, Lemma 2 implies $\theta'' \leq \theta^{**} < \theta^* \leq \theta'$, which is a contradiction, since by definition $\theta' \leq \theta''$. Therefore, there exists no robust equilibrium with $r^* \in (\widetilde{r},\overline{r}]$. On the other hand, in any robust equilibrium where $r^* \in (\underline{r},\widetilde{r}]$ is played, necessarily $\theta^* \leq \theta' \leq \theta'' \leq \theta^{**}$.

Next, we show, by iterated deletion of strictly dominated strategies, that in any robust equilibrium in which $r^* \in (\underline{r},\widetilde{r}]$ is played, the status quo is abandoned if and only if $\theta < \theta^*$, which in turn implies that $\theta' = \theta^*$.

Lemma 3. $D(\theta) = 1$ for all $\theta < \theta^*$, $D(\theta) = 0$ for all $\theta > \theta^*$, and $\theta' = \theta^*$.

Proof. Given an arbitrary equilibrium policy $r(\theta)$, we consider the continuation game that follows $r = \underline{r}$ and construct the iterated deletion mapping $\mathcal{T} : [\underline{\theta}, \theta'] \to [\underline{\theta}, \theta']$ as follows. Take any $\widetilde{\theta} \in [\underline{\theta}, \theta']$ and let

$$G(x;\widetilde{\theta}) \equiv \frac{1-\Psi\left(\frac{x-\widetilde{\theta}}{\varepsilon}\right)}{1-\Psi\left(\frac{x-\theta^*}{\varepsilon}\right)+p(x)+\Psi\left(\frac{x-\theta''}{\varepsilon}\right)}.$$

$G(x;\widetilde{\theta})$ thus represents the probability that $\theta < \widetilde{\theta}$, conditional on x and $\underline{r}$. Note that $\lim_{x\to-\infty} p(x) = \lim_{x\to+\infty} p(x) = 0$ and therefore $\lim_{x\to-\infty} G(x;\widetilde{\theta}) = 1$ and $\lim_{x\to+\infty} G(x;\widetilde{\theta}) = 0$. It follows that, for every $\widetilde{\theta}$, there is at least one solution to the equation $G(x;\widetilde{\theta}) = \underline{r}$. Then let $\widetilde{x} = \widetilde{x}(\widetilde{\theta})$ be the lowest solution to this equation, i.e. $\widetilde{x}(\widetilde{\theta}) \equiv \min\{x \mid G(x;\widetilde{\theta}) = \underline{r}\}$, and define $\widetilde{\widetilde{\theta}} = \widetilde{\widetilde{\theta}}(\widetilde{\theta})$ as the unique solution to $\widetilde{\widetilde{\theta}} = \Psi\left(\frac{\widetilde{x}-\widetilde{\widetilde{\theta}}}{\varepsilon}\right)$. Note that $G(x;\widetilde{\theta}) > \underline{r}$ for every $x < \widetilde{x}$ and $\Psi\left(\frac{\widetilde{x}-\theta}{\varepsilon}\right) > \theta$ for every $\theta < \widetilde{\widetilde{\theta}}$. That is, if private agents expect the status quo to be abandoned for all $\theta < \widetilde{\theta}$, necessarily they find it optimal to attack for all $x < \widetilde{x}$, which in turn implies that a regime change necessarily occurs for any $\theta < \widetilde{\widetilde{\theta}}$, unless $r(\theta) = r^*$ rather than $\underline{r}$, which happens in equilibrium only if $\theta \geq \theta'$. The iterated deletion operator $\mathcal{T}$ is thus defined by

$$\mathcal{T}\left(\widetilde{\theta}\right) \equiv \min\left\{\theta', \widetilde{\widetilde{\theta}}\right\}.$$

Observe that G is strictly increasing in $\widetilde{\theta}$, implying that $\widetilde{x}$ and therefore $\widetilde{\widetilde{\theta}}$ are also strictly increasing in $\widetilde{\theta}$.[22] We conclude that the mapping $\mathcal{T}$ is weakly increasing for all $\widetilde{\theta} \in [\underline{\theta}, \theta']$. Obviously, $\mathcal{T}$ is also bounded above by θ'. Finally, note that $\theta^* \leq \theta' \leq \theta^{**} < \infty$ and $\theta^* \leq \theta_{MS}$, but so far we have ruled out neither $\theta' \leq \theta_{MS}$, nor $\theta' > \theta_{MS}$.

Next, we compare $\mathcal{T}$ with the iterated deletion operator of the Morris-Shin game without signaling (or, equivalently, of the continuation game at $\underline{r}$ when the pooling equilibrium is played). This operator, $\mathcal{F} : [\underline{\theta}, \overline{\theta}] \to [\underline{\theta}, \overline{\theta}]$, is defined by $\mathcal{F}(\widehat{\theta}) = \widehat{\widehat{\theta}}$, where $\widehat{x} = \widehat{x}(\widehat{\theta})$ and $\widehat{\widehat{\theta}} = \widehat{\widehat{\theta}}(\widehat{\theta})$ are the unique solutions to $1-\Psi\left(\frac{\widehat{x}-\widehat{\theta}}{\varepsilon}\right) = \underline{r}$ and $\widehat{\widehat{\theta}} = \Psi\left(\frac{\widehat{x}-\widehat{\widehat{\theta}}}{\varepsilon}\right)$. $\mathcal{F}$ has a unique fixed point at $\widehat{\theta} = \widehat{\widehat{\theta}} = \theta_{MS}$, and satisfies $\widehat{\theta} < \mathcal{F}(\widehat{\theta}) < \theta_{MS}$ whenever $\widehat{\theta} \in [\underline{\theta}, \theta_{MS})$ and $\widehat{\theta} > \mathcal{F}(\widehat{\theta}) > \theta_{MS}$ whenever $\widehat{\theta} \in (\theta_{MS}, \overline{\theta}]$. Moreover, since $G(x;\theta) > 1-\Psi\left(\frac{x-\theta}{\varepsilon}\right)$ for all x and θ, we have $\widetilde{x}(\theta) > \widehat{x}(\theta)$ and therefore $\widetilde{\widetilde{\theta}}(\theta) > \widehat{\widehat{\theta}}(\theta)$. We conclude that, for any $\theta \in [\underline{\theta}, \theta']$, either $\widetilde{\widetilde{\theta}}(\theta) \geq \theta'$, in which case $\mathcal{T}(\theta) = \theta'$, or $\widetilde{\widetilde{\theta}}(\theta) < \theta'$, in which case $\mathcal{T}(\theta) > \mathcal{F}(\theta)$.

---

[22] In general, $\widetilde{x}$ and therefore $\widetilde{\widetilde{\theta}}$ need not be continuous in $\widetilde{\theta}$. Continuity is ensured when Ψ satisfies the MLRP, in which case G is strictly decreasing in x, implying that $G(x;\widetilde{\theta}) = \underline{r}$ has a unique solution and this solution is continuously increasing in $\widetilde{\theta}$. All we need, however, is monotonicity of the lowest solution, which is true for any Ψ.

Finally, we consider the sequence $\{\theta^k\}_{k=1}^{\infty}$, where $\theta^1 = \underline{\theta}$ and $\theta^{k+1} = \mathcal{T}(\theta^k)$ for all $k \geq 1$. This sequence represents iterated deletion of strictly dominated strategies starting from $\underline{\theta}$. Since this sequence is monotonic and bounded above by θ', it necessarily converges to some limit $\theta^{\infty} \leq \theta'$. This limit must be a fixed point of $\mathcal{T}$, that is, $\theta^{\infty} = \mathcal{T}(\theta^{\infty})$. We first prove that either $\theta^{\infty} = \theta'$, or $\theta^{\infty} > \theta_{MS}$. Suppose $\theta^{\infty} < \theta'$. This can be true only if $\widetilde{\widetilde{\theta}}(\theta^{\infty}) < \theta'$. If it were the case that $\theta^{\infty} \leq \theta_{MS}$, we would then have $\mathcal{T}(\theta^{\infty}) > \mathcal{F}(\theta^{\infty}) \geq \theta^{\infty}$, which contradicts the assumption that θ^{∞} is a fixed point for $\mathcal{T}(\cdot)$. Therefore, $\theta^{\infty} = \theta'$ whenever $\theta' \leq \theta_{MS}$, whereas $\theta^{\infty} > \theta_{MS}$ whenever $\theta' > \theta_{MS}$. We next prove that $\theta^{\infty} = \theta' = \theta^*$. Suppose that $\theta' > \theta^*$. If $\theta' \leq \theta_{MS}$, then $\theta^{\infty} = \theta' > \theta^*$ and all $\theta \in (\theta^*, \theta')$ would abandon the status quo in equilibrium. If instead $\theta' > \theta_{MS}$, then $\theta^{\infty} > \theta_{MS} \geq \theta^*$ and again all $\theta \in (\theta^*, \theta^{\infty})$ would abandon the status quo in equilibrium. But either case is impossible as any $\theta > \theta^*$ can ensure no regime change and a positive payoff by setting r^*. Therefore, it is necessarily the case that $\theta' = \theta^*$ and since $\theta^* \leq \theta_{MS}$ then $\theta^{\infty} = \theta' = \theta^*$. It follows that $D(\theta) = 1$ for all $\theta < \theta^*$ and $D(\theta) = 0$ for all $\theta > \theta^*$. □

Combining the above three lemmas, we conclude that any robust equilibrium belongs necessarily to either (a) or (b) in Theorem 1. which completes the proof of the theorem. ■

Proof of Remark on Monotone Strategies/MLRP The proof is in two steps. We first prove that if $a(x, \underline{r})$ is monotonic (decreasing) in x, the size of an attack $A(\theta, \underline{r})$ is decreasing in the fundamentals θ which in turns implies that the policy maker sets r^* for all $\theta \in [\theta^*, \theta^{**}]$. Next, we prove that if Ψ satisfies the MLRP, i.e. if ψ'/ψ is monotonic decreasing, then $a(x, \underline{r})$ is necessarily monotonic in x.

Step 1. If in equilibrium $a(x, \underline{r})$ is monotonic in x, then $\theta'' = \theta^{**}$, and $r(\theta) = r^*$ for all $\theta \in [\theta^*, \theta^{**}]$.

Proof. If $a(x, \underline{r})$ is monotonic in x, then necessarily $A(\theta, \underline{r})$ is decreasing in θ. From the indifference condition $A(\theta'', \underline{r}) = C(r^*)$, it follows that all $\theta \in [\theta^*, \theta'']$ necessarily set r^*. In this case, the probability of regime change conditional on x and $\underline{r}$ is given by

$$\mu\left(D = 1|x, \underline{r}\right) = \mu(\theta \leq \theta^*|x, \underline{r}) = \frac{1 - \Psi\left(\frac{x - \theta^*}{\varepsilon}\right)}{1 - \Psi\left(\frac{x - \theta^*}{\varepsilon}\right) + \Psi\left(\frac{x - \theta''}{\varepsilon}\right)},$$

which is decreasing in x. Since $\lim_{x \to -\infty} \mu\left(D = 1|x, \underline{r}\right) = 1$ and $\lim_{x \to +\infty} \mu\left(D = 1|x, \underline{r}\right) = 0$, there exists a unique $\widehat{x} = \widehat{x}\left(\theta^*, \theta''\right)$ such that $\mu\left(D = 1|\widehat{x}, \underline{r}\right) = \underline{r}$ and thus private agents follow a threshold strategy with $a(x, \underline{r}) = 1$ if $x < \widehat{x}$ and $a(x, \underline{r}) = 0$ if $x \geq \widehat{x}$ implying that $A(\theta, \underline{r}) = \Psi\left(\frac{\widehat{x}(\theta^*, \theta'') - \theta}{\varepsilon}\right)$. Thus, θ'' solves $\Psi\left(\frac{\widehat{x}(\theta^*, \theta'') - \theta''}{\varepsilon}\right) = C(r^*)$, and from Proposition (3), $\widehat{x}\left(\theta^*, \theta^{**}\right) - \theta^{**} = \widehat{x}\left(\theta^*, \theta''\right) - \theta''$.

From the monotonicity of $\widehat{x}(\theta^*, \theta) - \theta$ established in Theorem (2) – Lemma 2 – it follows that $\theta'' = \theta^{**}$. □

Step 2. If Ψ satisfies the MLRP, then $\mu(\theta \leq \theta^*|x, \underline{r})$ and thus $a(x, \underline{r})$ are necessarily decreasing in x.

Proof. Let $I(\theta)$ be the probability that θ sets $\underline{r}$. Using the fact that $I(\theta) = 1$ for all $\theta \leq \theta^*$, the probability of regime change conditional on x and $\underline{r}$ can be written as

$$\mu(\theta \leq \theta^*|x, \underline{r}) = \frac{1 - \Psi\left(\frac{x-\theta^*}{\varepsilon}\right)}{1 - \Psi\left(\frac{x-\theta^*}{\varepsilon}\right) + \int_{\theta^*}^{\infty} \frac{1}{\varepsilon}\psi\left(\frac{x-\theta}{\varepsilon}\right) I(\theta)\, d\theta}.$$

Using

$$M(x) \equiv \frac{\mu(\theta \leq \theta^*|x, \underline{r})}{1 - \mu(\theta \leq \theta^*|x, \underline{r})} = \frac{1 - \Psi\left(\frac{x-\theta^*}{\varepsilon}\right)}{\int_{\theta^*}^{\infty} \frac{1}{\varepsilon}\psi\left(\frac{x-\theta}{\varepsilon}\right) I(\theta)\, d\theta},$$

we have that

$$\frac{dM}{dx} = \frac{-\frac{1}{\varepsilon}\psi\left(\frac{x-\theta^*}{\varepsilon}\right)}{\int_{\theta^*}^{\infty} \frac{1}{\varepsilon}\psi\left(\frac{x-\theta}{\varepsilon}\right) I(\theta)\, d\theta} - \frac{\left[1 - \Psi\left(\frac{x-\theta^*}{\varepsilon}\right)\right] \frac{d}{dx}\left(\int_{\theta^*}^{\infty} \frac{1}{\varepsilon}\psi\left(\frac{x-\theta}{\varepsilon}\right) I(\theta)\, d\theta\right)}{\left[\int_{\theta^*}^{\infty} \frac{1}{\varepsilon}\psi\left(\frac{x-\theta}{\varepsilon}\right) I(\theta)\, d\theta\right]^2}.$$

It follows that $d\mu(\theta \leq \theta^*|x, \underline{r})/dx \leq 0$ if and only if

$$\frac{\int_{-\infty}^{\theta^*} \frac{1}{\varepsilon^2}\psi'\left(\frac{x-\theta}{\varepsilon}\right) d\theta}{\int_{-\infty}^{\theta^*} \frac{1}{\varepsilon}\psi\left(\frac{x-\theta}{\varepsilon}\right) d\theta} - \frac{\int_{\theta^*}^{\infty} \frac{1}{\varepsilon^2}\psi'\left(\frac{x-\theta}{\varepsilon}\right) I(\theta)\, d\theta}{\int_{\theta^*}^{\infty} \frac{1}{\varepsilon}\psi\left(\frac{x-\theta}{\varepsilon}\right) I(\theta)\, d\theta} \leq 0.$$

Using the fact that $I(\theta) = 1$ for all $\theta \leq \theta^*$, the above is equivalent to

$$\mathbb{E}_\theta\left[\frac{\frac{1}{\varepsilon}\psi'\left(\frac{x-\theta}{\varepsilon}\right)}{\psi\left(\frac{x-\theta}{\varepsilon}\right)} \middle| \theta \leq \theta^*,\ x,\ \underline{r}\right] - \mathbb{E}_\theta\left[\frac{\frac{1}{\varepsilon}\psi'\left(\frac{x-\theta}{\varepsilon}\right)}{\psi\left(\frac{x-\theta}{\varepsilon}\right)} \middle| \theta > \theta^*,\ x,\ \underline{r}\right] \leq 0,$$

which holds true if ψ'/ψ is monotonic decreasing.

From the monotonicity of $\mu(\theta \leq \theta^*|x, \underline{r})$, it follows immediately that $a(x, \underline{r})$ is monotonic with cutoff x' given by $\mu(\theta \leq \theta^*|x', \underline{r}) = \underline{r}$.

Proof of Proposition 4. For $\theta < \underline{\theta}$, it is dominant for the policy maker to set $\underline{r}$ and abandon the status quo and for private agents to attack. Similarly, for $\theta > \overline{\theta}$, the status quo is never abandoned, private agents do not attack, and there is no need to undertake any costly policy measure. Finally, take any $\theta \in [\underline{\theta}, \overline{\theta}]$. The continuation game following any level of the policy r is a coordination game with two (extreme) continuation equilibria, no attack and full attack. Let $r(\theta)$ be the minimal r

for which private agents coordinate on the no-attack continuation equilibrium, i.e. they attack if and only if $r < r(\theta)$. Clearly, it is optimal for the policy maker to set $r(\theta) > \underline{r}$ if and only if $V(\theta, 0) - C(r(\theta)) \geq 0$, or equivalently $C(r(\theta)) \leq \theta$.

Proof of Proposition 5. Take any random variable r^* with compact support $\mathcal{R}^* \subseteq (\underline{r}, \widetilde{r})$ and distribution Φ. We want to show that for ε sufficiently small, there exist thresholds $x^*, \theta^* \in (\underline{\theta}, \theta_{MS})$, and $\theta^{**} \in (\theta^*, \overline{\theta})$, a system of beliefs μ, and a robust equilibrium such that: (i) The policy maker sets $r(\theta) = \underline{r}$ for $\theta \notin [\theta^*, \theta^{**}]$, and $r(\theta) \in \mathcal{R}^*$ with $r(\theta)$ non-decreasing in θ for $\theta \in [\theta^*, \theta^{**}]$. (ii) Whenever $r < \underline{r}^*$, private agents attack if and only if $x < x^*$; whenever $r \in [\underline{r}^*, r^*)$, they attack if and only if $x < \overline{x}$; and whenever $r \geq r^*$, they attack if and only if $x < \underline{x}$. (iii) The status quo is abandoned with probability $D(\theta) = 1$ for $\theta < \theta^*$, with probability $D(\theta) < 1$ and non-increasing in θ for $\theta \in [\theta^*, \theta^{**}]$, and with probability $D(\theta) = 0$ for $\theta > \theta^{**}$.

The proof is in five steps: Steps 1 and 2 characterize the thresholds θ^*, θ^{**}, and x^*; Step 3 examines the behavior of the policy maker; Step 4 examines the beliefs and the strategy of the private agents; Step 5 establishes robustness in the sense of Definition 2.

Step 1. Let $\underline{r}^* \equiv \min \mathcal{R}^* > \underline{r}$ and $\overline{r}^* \equiv \max \mathcal{R}^* < \widetilde{r}$. Define

$$\widehat{U}(\theta) \equiv \begin{cases} \max\limits_{r \in [\underline{r}^*, \overline{r}^*]} \{-C(r)\} & \text{if } \theta < \underline{\theta}, \\ \max\limits_{r \in [\underline{r}^*, \overline{r}^*]} \{\theta \Phi(r) - C(r)\} & \text{if } \theta \in [\underline{\theta}, \overline{\theta}], \\ \max\limits_{r \in [\underline{r}^*, \overline{r}^*]} \left\{\theta - (1 - \Phi(r)) \Psi\left(\frac{\overline{x} - \theta}{\varepsilon}\right) - C(r)\right\} & \text{if } \theta > \overline{\theta}, \end{cases}$$

and $\widehat{r}(\theta)$ as the corresponding arg max. Note that maximizing over $[\underline{r}^*, \overline{r}^*]$ is equivalent to maximizing over $\mathcal{R}^*$, since necessarily $\widehat{r}(\theta) \in \mathcal{R}^*$. Note also that $\widehat{U}(\theta)$ is non-decreasing for all θ, and strictly increasing whenever $\widehat{U}(\theta) > -C(\underline{r}^*)$. Moreover, $\widehat{U}(\underline{\theta}) = -C(\underline{r}^*) < 0$ and $\widehat{U}(\theta_{MS}) \geq \theta_{MS} - C(\overline{r}^*) > \theta_{MS} - C(\widetilde{r}) = 0$. Therefore, there exists a unique $\theta^* \in (\underline{\theta}, \theta_{MS})$ such that $\widehat{U}(\theta^*) = 0$, $\widehat{U}(\theta) < 0$ for all $\theta < \theta^*$, and $\widehat{U}(\theta) > 0$ for all $\theta > \theta^*$.

Step 2. For any $\theta \geq \theta^*$, define the functions $\widehat{x}(\theta; \theta^*)$ and $v(\theta; \theta^*)$ as follows

$$\frac{1 - \Psi\left(\frac{\widehat{x} - \theta^*}{\varepsilon}\right)}{1 - \Psi\left(\frac{\widehat{x} - \theta^*}{\varepsilon}\right) + \Psi\left(\frac{\widehat{x} - \theta}{\varepsilon}\right)} = \underline{r} \quad \text{and} \quad v(\theta; \theta^*) = \theta - \Psi\left(\frac{\widehat{x}(\theta; \theta^*) - \theta}{\varepsilon}\right).$$

Note that $d\widehat{x}(\theta; \theta^*)/d\theta \in (0, 1)$ and therefore $dv(\theta; \theta^*)/d\theta > 1$. Note also that for any $\theta > \theta^*$, $\Psi\left(\frac{\widehat{x}(\theta;\theta^*) - \theta}{\varepsilon}\right) \to 0$ as $\varepsilon \to 0$. To see this, suppose instead that $\lim_{\varepsilon \to 0} \Psi\left(\frac{\widehat{x}(\theta;\theta^*) - \theta}{\varepsilon}\right) = \omega$ for some

$\omega > 0$. This can be true only if $\lim_{\varepsilon\to 0} \widehat{x}(\theta;\theta^*) \geq \theta$, in which case $\theta > \theta^*$ implies $\lim_{\varepsilon\to 0} \widehat{x}(\theta;\theta^*) > \theta^*$ and therefore $\lim_{\varepsilon\to 0} \Psi\left(\frac{\widehat{x}(\theta;\theta^*)-\theta^*}{\varepsilon}\right) = 1$. But then

$$\lim_{\varepsilon\to 0}\left\{\frac{1-\Psi\left(\frac{\widehat{x}(\theta;\theta^*)-\theta^*}{\varepsilon}\right)}{1-\Psi\left(\frac{\widehat{x}(\theta;\theta^*)-\theta^*}{\varepsilon}\right)+\Psi\left(\frac{\widehat{x}(\theta;\theta^*)-\theta}{\varepsilon}\right)}\right\} = \frac{0}{0+\omega} = 0 < \underline{r},$$

which is a contradiction. Consider now the function $g(\theta) = v(\theta;\theta^*) - \widehat{U}(\theta)$. From the envelope theorem, $d\widehat{U}(\theta)/d\theta$ is bounded above by 1 for any $\theta \leq \overline{\theta}$ and hence $g(\theta)$ is strictly increasing in θ over this range. Moreover, by definition of $\widehat{x}(\theta;\theta^*)$, we have that $\Psi\left(\frac{\widehat{x}(\theta^*;\theta^*)-\theta^*}{\varepsilon}\right) = 1 - \underline{r} = \theta_{MS}$ and therefore $v(\theta^*;\theta^*) = \theta^* - \theta_{MS}$. Since $\theta^* < \theta_{MS}$ and $\widehat{U}(\theta^*) = 0$, we conclude that $g(\theta^*) < 0$. Next, note that for any $\theta \leq \overline{\theta}$, $\widehat{r}(\cdot)$ is independent of ε, whereas $\Psi\left(\frac{\widehat{x}(\theta;\theta^*)-\theta}{\varepsilon}\right) \to 0$ as $\varepsilon \to 0$ for every $\theta > \theta^*$. (See the argument above). Since $\overline{\theta} > \theta^*$ and $C(\widehat{r}(\overline{\theta})) > 0$, it follows that there exists $\overline{\varepsilon} > 0$ such that $\Psi\left(\frac{\widehat{x}(\overline{\theta};\theta^*)-\overline{\theta}}{\varepsilon}\right) < C(\widehat{r}(\overline{\theta}))$ whenever $\varepsilon < \overline{\varepsilon}$. But then $v(\overline{\theta};\theta^*) > \overline{\theta} - C(\widehat{r}(\overline{\theta})) \geq \widehat{U}(\overline{\theta})$ and therefore $g(\overline{\theta}) > 0$. We conclude that, for $\varepsilon < \overline{\varepsilon}$, there exists a unique $\theta^{**} \in (\theta^*, \overline{\theta})$ such that $g(\theta^{**}) = 0$, $g(\theta) < 0$ for $\theta < \theta^{**}$, and $g(\theta) > 0$ for $\theta > \theta^{**}$. Moreover, note that, as $\varepsilon \to 0$, $v(\theta;\theta^*) \to \theta > \widehat{U}(\theta)$ for every $\theta > \theta^*$; it follows that $\theta^{**} \to \theta^*$ as $\varepsilon \to 0$. Next, let

$$x^* \equiv \widehat{x}(\theta^{**};\theta^*) \qquad \text{and} \qquad \widetilde{U}(\theta;\theta^*) \equiv \theta - \Psi\left(\frac{x^*-\theta}{\varepsilon}\right).$$

Compare now $\widetilde{U}(\theta;\theta^*)$ with $v(\theta;\theta^*)$. That $\widehat{x}(\theta^{**};\theta^*) = x^*$ implies $\widetilde{U}(\theta^{**};\theta^*) = v(\theta^{**};\theta^*)$, while the fact that $\widehat{x}(\theta;\theta^*)$ is increasing in θ implies $\widetilde{U}(\theta;\theta^*) < v(\theta;\theta^*)$ for all $\theta < \theta^{**}$ and $\widetilde{U}(\theta;\theta^*) > v(\theta;\theta^*)$ for all $\theta > \theta^{**}$. Combining this result with the properties of the function $g(\theta)$, we conclude that $\widetilde{U}(\theta;\theta^*) < v(\theta;\theta^*) < \widehat{U}(\theta)$ for all $\theta < \theta^{**}$, $\widetilde{U}(\theta;\theta^*) = v(\theta;\theta^*) = \widehat{U}(\theta)$ at $\theta = \theta^{**}$, and $\widetilde{U}(\theta;\theta^*) > v(\theta;\theta^*) > \widehat{U}(\theta)$ for all $\theta > \theta^{**}$. Finally, let $\widehat{\theta}$ be the unique solution to $\widetilde{U}(\widehat{\theta};\theta^*) = 0$ and note that $\widehat{\theta} \in (\theta^*, \theta^{**})$, since $\widetilde{U}(\theta;\theta^*)$ is increasing in θ and $\widetilde{U}(\theta^{**};\theta^*) > 0 > \widetilde{U}(\theta^*;\theta^*)$.

Step 3. Consider now the behavior of the policy maker. Given the strategy of the private agents, the policy maker prefers $\underline{r}$ to any $r < \underline{r}^*$, and $\overline{r}^*$ to any $r > \overline{r}^*$. Furthermore, by definition, $\widehat{r}(\theta)$ dominates any $r \in [\underline{r}^*, \overline{r}^*]$. We thus need to compare only the payoff from playing $\widehat{r}(\theta)$ with that from playing $\underline{r}$. Playing $\widehat{r}(\theta)$ yields $\widehat{U}(\theta)$, while playing $\underline{r}$ yields $\underline{U}(\theta) = \max\{0, \widetilde{U}(\theta;\theta^*)\}$. Note that $\underline{U}(\theta) = 0$ if $\theta \leq \widehat{\theta}$ and $\underline{U}(\theta) = \widetilde{U}(\theta) > 0$ if $\theta > \widehat{\theta} \in (\theta^*, \theta^{**})$. It follows that $\widehat{U}(\theta) < 0 = \underline{U}(\theta)$ for all $\theta < \theta^*$, $\widehat{U}(\theta) > 0 = \underline{U}(\theta)$ for all $\theta \in (\theta^*, \widehat{\theta}]$, $\widehat{U}(\theta) > \widetilde{U}(\theta) = \underline{U}(\theta) > 0$ for all $\theta \in (\widehat{\theta}, \theta^{**})$, and $\underline{U}(\theta) = \widetilde{U}(\theta) > \widehat{U}(\theta)$ for all $\theta > \theta^{**}$. Therefore, it is indeed optimal for the policy maker to play $\widehat{r}(\theta)$ whenever $\theta \in [\theta^*, \theta^{**}]$ and $\underline{r}$ otherwise. The resulting probability of regime change is $D(\theta) = 1$

for $\theta < \theta^*$, $D(\theta) = 1 - \Phi(\widehat{r}(\theta)) \in [0,1)$ for $\theta \in [\theta^*, \theta^{**}]$, and $D(\theta) = 0$ for $\theta > \theta^{**}$. Since $\widehat{r}$ is non-decreasing, D is non-increasing in the range $[\theta^*, \theta^{**}]$.

Step 4. Consider next the behavior of the private agents. For any $r < \underline{r}^*$, the status quo is abandoned if and only if $\theta < \widehat{\theta}$. In equilibrium, at $r = \underline{r}$, the probability of regime change is given by

$$\mu(\theta < \widehat{\theta}|x, \underline{r}) = \mu(\theta < \theta^*|x, \underline{r}) = \frac{1 - \Psi\left(\frac{x-\theta^*}{\varepsilon}\right)}{1 - \Psi\left(\frac{x-\theta^*}{\varepsilon}\right) + \Psi\left(\frac{x-\theta^{**}}{\varepsilon}\right)}.$$

By construction, x^* solves $\mu(\theta < \widehat{\theta}|x^*, \underline{r}) = \underline{r}$, and since $\mu(\theta < \widehat{\theta}|x, \underline{r})$ is decreasing in x, attacking the regime if and only if $x < x^*$ is indeed optimal. For any out-of-equilibrium $r < \underline{r}^*$, we consider beliefs μ such that $\mu(\theta < \widehat{\theta}|x, r)$ is non-increasing in x and $\mu(\theta < \widehat{\theta}|x, r) = r$ at $x = x^*$. For any $r \geq \underline{r}^*$, we let $\mu(\theta \in (\underline{\theta}, \overline{\theta})|x, r) = 1$ for all $x \in [\underline{x}, \overline{x}]$, $\mu(\theta < \underline{\theta}|x, r) = 1$ for $x < \underline{x}$, and $\mu(\theta > \overline{\theta}|x, r) = 1$ for $x > \overline{x}$. In particular, for any $r \in r(\Theta)$ with $r \geq \underline{r}^*$, $\mu(\theta \in \theta^{-1}(r)|x, r) = 1$ for all x such that $\Theta(x) \cap \theta^{-1}(r) \neq \varnothing$, where $\theta^{-1}(r) = \{\theta : r(\theta) = r\}$. In addition, for any out-of-equilibrium r, if r is not dominated in equilibrium by $r(\theta)$ for some $\theta \in \Theta(x)$, then we further restrict μ to satisfy $\mu(\theta|x, r) = 0$ for all $\theta \in \Theta(x)$ such that $\theta - C(r) < U(\theta)$. These beliefs satisfy (4), as well as $\mu \in \mathcal{M}(r)$; and given these beliefs, the strategy of the agents is sequentially optimal for any r and x.

Step 5. Finally, consider robustness. Assume first that Ψ has bounded support, let ε be sufficiently small, and construct the corresponding sunspot equilibrium as above. Let $A^* \equiv A(\theta^*, \underline{r}) = \Psi\left(\frac{x^*-\theta^*}{\varepsilon}\right)$ and $A^{**} \equiv A(\theta^{**}, \underline{r}) = \Psi\left(\frac{x^*-\theta^{**}}{\varepsilon}\right)$ and note that $0 < A^{**} < A^* < 1$ as $\mu(\theta < \theta^*|x, \underline{r}) = \underline{r} \in (0,1)$. Similarly, let $\xi^* \equiv \Psi^{-1}(A^*)$ and $\xi^{**} \equiv \Psi^{-1}(A^{**})$ and note that $\xi^{**} < \xi^*$. Hence, one can always find a $k > 0$ such that $0 < \Psi(\xi^{**} - k) < \Psi(\xi^* + k) < 1$, and a c.d.f. F with unbounded support such that $F(\xi) = \Psi(\xi)$ for all $\xi \in [\xi^{**} - k, \xi^* + k]$, $F(\xi) > \Psi(\xi)$ for all $\xi < \xi^{**} - k$, and $F(\xi) < \Psi(\xi)$ for all $\xi > \xi^* + k$. We argue that the same $r(\theta)$ and $D(\theta)$ that is sustained when the noise distribution is Ψ, can also be sustained when the distribution is F. Indeed, consider the functions $\widehat{U}, \widehat{r}, \widehat{x}, v, g, \widetilde{U}$, and $\underline{U}$ defined above, where Ψ is replaced with F. Note that $\widehat{U}(\theta)$ and $\widehat{r}(\theta)$ are independent of the noise distribution for all $\theta \leq \overline{\theta}$. It follows that the same θ^* continues to solve $\widehat{U}(\theta^*) = 0$. By construction of F, the functions $\widehat{x}(\theta; \theta^*)$, $v(\theta; \theta^*)$, and $g(\theta)$ remain the same in a neighborhood of θ^{**}. It follows that the same θ^{**} and x^* continue to solve $g(\theta^{**}) = 0$ and $x^* = \widehat{x}(\theta^{**}; \theta^*)$. But then $\widetilde{U}(\theta; \theta^*)$ and $\underline{U}(\theta)$ also remain the same in a neighborhood of θ^{**}. Since $\widehat{U}(\theta)$ is also the same for all $\theta < \overline{\theta}$, and $\theta^{**} < \overline{\theta}$, we infer $\underline{U}(\theta^{**}) = \widetilde{U}(\theta^{**}; \theta^*) = \widehat{U}(\theta^{**})$. Like in Step 3 above, the latter ensures that the optimal policy is $r(\theta) = \widehat{r}(\theta)$ for all $\theta \in [\theta^*, \theta^{**}]$

and $r(\theta) = \underline{r}$ otherwise. Since $\widehat{r}(\theta)$ is also the same for all $\theta \in [\theta^*, \theta^{**}]$, this proves that the same policy $r(\theta)$ can be sustained with either Ψ or F. Similarly, all $\theta < \theta^*$ continue to abandon the status quo with certainty and all $\theta > \theta^{**}$ continue to maintain the status quo with certainty, while for $\theta \in [\theta^*, \theta^{**}]$ the probability of regime change is $D(\theta) = \Phi(\widehat{r}(\theta))$. Therefore, the same probability of regime change$D(\theta)$ can also be sustained with either Ψ or F. A symmetric argument applies if Ψ is unbounded. With ξ^* and ξ^{**} defined as above, take an arbitrary $k > 0$ and construct F so that F has bounded support $[\xi^{**} - 2k, \xi^* + 2k]$ and satisfies $F(\xi) = \Psi(\xi)$ for all $\xi \in [\xi^{**} - k, \xi^* + k]$, $F(\xi) < \Psi(\xi)$ for all $\xi < \xi^{**} - k$, and $F(\xi) > \Psi(\xi)$ for all $\xi > \xi^{**} - k$. It follows that F sustains the same $r(\theta)$ and $D(\theta)$ as Ψ. We conclude that the sunspot equilibria of Proposition 5 are robust. ■

References

[1] Albanesi, Stefania, V.V. Chari, and Lawrence Christiano, (2002), "Expectation Traps and Monetary Policy," NBER working paper 8912.

[2] Angeletos, George-Marios (2003), "Idiosyncratic Investment Risk in a Neoclassical Growth Economy," NBER working paper.

[3] Atkeson, Andrew (2000), "Discussion on Morris and Shin," *NBER Macro Annual.*

[4] Azariadis, Costas (1981), "Self-fulfilling Prophecies," *Journal of Economic Theory* 25, 380-96.

[5] Battaglini, Marco, and Roland Benabou, (2003), "Trust, Coordination, and the Industrial Organization of Political Activism," *Journal of the European Economic Association* 1.

[6] Benhabib, Jess, and Roger Farmer (1994), "Indeterminacy and Increasing Returns," *Journal of Economic Theory* 63, 19-41.

[7] Bernanke, Ben, and Mark Gertler (1989), "Agency Costs, Net Worth, and Business Fluctuations," *American Economic Review* 79, 14-31.

[8] Bryant, John (1983), "A Simple Rational Expectations Keynes-Type Model," *Quarterly Journal of Economics* 98, 3, 525-28.

[9] Calvo, G. (1988), "Servicing the Public Debt: the Role of Expectations." *American Economic Review,* 78(4), 647-661.

[10] Carlsson, Hans, and Eric van Damme (1993), "Global Games and Equilibrium Selection," *Econometrica* 61, 5, 989-1018.

[11] Chamley, Christophe (1999), "Coordinating Regime Switches," *Quarterly Journal of Economics* 114, 3, 869-905.

[12] Chari, V.V., Lawrence Christiano, and Martin Eichenbaum (1998), "Expectation Traps and Discretion," *Journal of Economic Theory* 81, 2, 462-92.

[13] Chari, V.V., and Patrick Kehoe (2003a), "Hot Money," Federal Reserve Bank of Minneapolis staff report 228.

[14] Chari, V.V., and Patrick Kehoe (2003b), "Financial Crises as Herds," Federal Reserve Bank of Minneapolis working paper 600.

[15] Cho, In-Koo, and David Kreps (1987), "Signaling Games and Stable Equilibria," *Quarterly Journal of Economics* 102, 2, 179-221.

[16] Cole, Harold, and Timothy Kehoe (1996), "Self-fulfilling Debt Crises," Federal Reserve Bank of Minneapolis staff report 211.

[17] Cooper, Russell (1998), *Coordination Games, Complementarities and Macroeconomics*, Cambridge University Press.

[18] Cooper, Russell, and Andrew John (1988), "Coordinating Coordination Failures in Keynesian Models," *Quarterly Journal of Economics* 103, 441–463.

[19] Diamond, Douglas, and Philip Dybvig (1983), "Bank Runs, Deposit Insurance, and Liquidity," *Journal of Political Economy* 91, 401-19.

[20] Diamond, Peter (1982), "Aggregate Demand Management in Search Equilibrium," *Journal of Political Economy* 90, 5, 881-94.

[21] Drazen, Alan (2001), "Interest Rate Defense Against Speculative Attack as a Signal: A Primer," University of Maryland mimeo.

[22] Farrell, J. and G. Saloner (1985), "Standardization, Compatibility, and Innovation," *Rand Journal of Economics* 16, 70-83.

[23] Feltovich, Nick, Rick Harbaugh, and Ted To (2002), "Too Cool for School? Signaling and Countersignaling," *Rand Journal of Economics* 33, 630-649.

[24] Freixas, Xavier, and Jean-Charles Rochet (1997), *Microeconomics of Banking*, Cambridge and London: MIT Press.

[25] Hellwig, Christian (2002), "Public Information, Private Information and the Multiplicity of Equilibria in Coordination Games," *Journal of Economic Theory*, 107, 191-222.

[26] Judd, Kenneth, (1985) "The Law of Large Numbers with a Continuum of IID Random Variables," *Journal of Economic Theory* 35, 19-25.

[27] Kajii A. and S. Morris (1997), "Common p-beliefs: The General Case," *Games and Economic Behavior*, 18, 73-82.

[28] Katz, M. and C. Shapiro (1985), "Network Externalities, Competition, and Compatibility," *American Economic Review* 75, 424-440.

[29] Katz, M. and C. Shapiro (1986), "Technology Adoption in the Presence of Network Externalities," *Journal of Political Economy* 94, 822-841.

[30] Kiyotaki, Nobuhiro (1988), "Multiple Expectational Equilibria under Monopolistic Competition," *Quarterly Journal of Economics* 103, 695–713.

[31] Kiyotaki, Nobuhiro, and John Moore (1997), "Credit Cycles," *Journal of Political Economy* 105, 211-248.

[32] Kraay, Aart (2003), "Do High Interest Rates Defend Currencies During Speculative Attacks?", *Journal of International Economics* 29, 297-321.

[33] Kuran, Timur (1987), "Preference Falsification, Policy Continuity and Collective Conservatism," *Economic Journal* 97, 642–665.

[34] Milgrom, Paul, and John Roberts (1990), "Rationalizability, Learning and Equilibrium in Games with Strategic Complementarities," *Econometrica*, 58, 1255-1277.

[35] Monderer, D. and D. Samet (1989), "Approximating Common Knowledge with Common Beliefs," *Games and Economic Behavior*, 1, 170-190.

[36] Morris, Stephen and Hyun Song Shin (1998), "Unique Equilibrium in a Model of Self-Fulfilling Currency Attacks", *American Economic Review*, 88, 3, 587-597.

[37] Morris, Stephen and Hyun Song Shin (2000), "Rethinking Multiple Equilibria in Macroeconomics," *NBER Macro Annual*.

[38] Morris, Stephen and Hyun Song Shin (2001), "Global Games - Theory and Applications," in *Advances in Economics and Econometrics*, 8th World Congress of the Econometric Society (M. Dewatripont, L. Hansen, and S. Turnovsky, eds.), Cambridge University Press, Cambridge, UK.

[39] Murphy, Kevin M., Andrei Shleifer, and Robert Vishny (1989), "Industrialization and the Big Push," *Journal of Political Economy* 97, 5, 1003-26.

[40] Obstfeld, Maurice (1986), "Rational and Self-Fulfilling Balance-of-Payments Crises", *American Economic Review* 76, 1, 72-81.

[41] Obstfeld, Maurice (1996), "Models of Currency Crises with Self-Fulfilling Features," *European Economic Review* 40, 3-5, 1037-47.

[42] Postlewaite, Andrew, and Xavier Vives (1987), "Bank Runs as an Equilibrium Phenomenon," *Journal of Political Economy* 95, 3, 485-91.

[43] Spence, Michael (1973), "Job Market Signaling," *Quarterly Journal of Economics* 87, 3, 355-374.

[44] Velasco, Andres (1996), "Fixed Exchange Rates: Credibility, Flexibility and Multiplicity," *European Economic Review* 40, 3-5, 1023-36.

[45] Vives, Xavier (1999), *Oligopoly Pricing: Old Ideas and New Tools*, MIT Press.

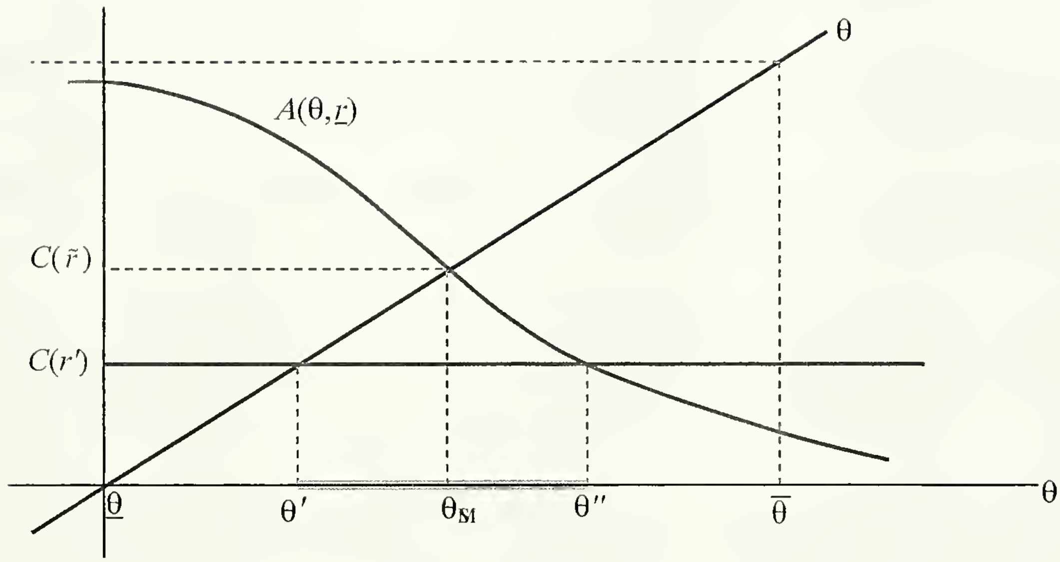

Figure 1

There exists an *inactive policy equilibrium* in which the policy maker sets $\underline{r}$ for all θ, the size of the attack is $A(\theta,\underline{r})$, and the status quo is abandoned if and only if $\theta < \theta_{SM}$. Any $r' \in (\underline{r}, \tilde{r}\,]$ is dominated in equilibrium by $\underline{r}$ if and only if $C(r')>\theta$ or $C(r')>A(\theta,\underline{r})$, that is, if and only if $\theta \notin [\theta',\theta'']$. It follows that, for any $\theta \in [\theta',\theta'']$, if the policy maker deviates from $\underline{r}$ to r', the market "learns" that $\theta \in [\theta',\theta'']$. Private agents then coordinate on the same behavior as when $r = \underline{r}$, thus eliminating any incentives for the policy maker to intervene.

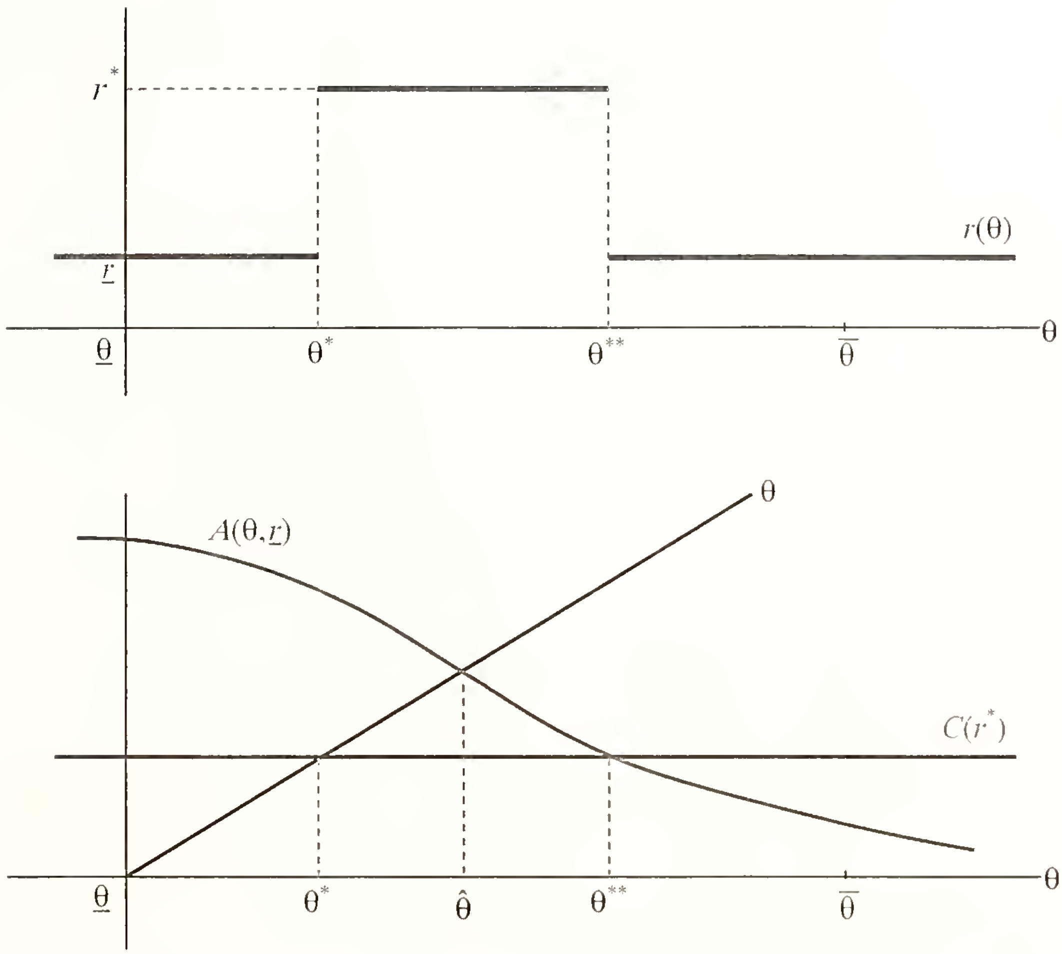

Figure 2

For each $r^* \in (\underline{r}, \tilde{r}]$, there exists a *two-threshold equilibrium* in which the policy is $r(\theta) = r^*$ if $\theta \in [\theta^*, \theta^{**}]$ and $r(\theta) = \underline{r}$ otherwise, and in which the status quo is abandoned if and only if $\theta < \theta^*$. When the policy maker raises the policy at r^*, the market "learns" that $\theta \in [\theta^*, \theta^{**}]$ and private agents coordinate on no attack. When instead the policy maker sets $\underline{r}$, private agents attack if and only if their signal is sufficiently low, in which case the size of the attack is decreasing in θ. It follows that it is optimal for the policy maker to raise the policy at r^* if and only if $C(r^*) \leq \theta$ and $C(r^*) \leq A(\theta, \underline{r})$, that is, if and only if $\theta \in [\theta^*, \theta^{**}]$.

3837 39

Date Due

Lib-26-67

MIT LIBRARIES
3 9080 02617 8241