Coordination and Policy Traps*

George-Marios Angeletos — MIT and NBER
Christian Hellwig — UCLA
Alessandro Pavan — Northwestern University

First draft: July 2002. This version: April 2005.

Abstract

This paper examines the ability of a policy maker to control equilibrium outcomes in a global coordination game; applications include currency attacks, bank runs, and debt crises. A unique equilibrium is known to survive when the policy is exogenously fixed. We show that, by conveying information, endogenous policy re-introduces multiple equilibria. Multiplicity obtains even in environments where the policy is observed with idiosyncratic noise. It is sustained by the agents coordinating on different interpretations of, and different reactions to, the same policy choices. The policy maker is thus trapped into a position where both the optimal policy and the coordination outcome are dictated by self-fulfilling market expectations.

JEL Classification Numbers: C7, D8, E5, E6, F3.

Key Words: global games, complementarities, signaling, self-fulfilling expectations, multiple equilibria, currency crises, regime change.

*The suggestions of the editor, Nancy Stokey, and two anonymous referees greatly improved the paper. For helpful comments, we also thank 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, Balazs Szentes, Jean Tirole, Muhamet Yildiz, Iván Werning, and seminar participants at AUEB, Berkeley, Carnegie-Mellon, Harvard, Iowa, Lausanne, LSE, Mannheim, MIT, Northwestern, Stanford, Stony Brook, Toulouse, UCLA, the FRBs in Minneapolis, Cleveland, and Boston, 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.

1 Introduction

Coordination failures are often invoked as justification for government intervention; they play a prominent role in currency attacks, debt crises, bank runs, investment crashes, and socio-political change.

A vast literature has captured the role of coordination in models that feature multiple equilibria.[1] Morris and Shin (1998, 2000, 2003), however, have argued that equilibrium multiplicity is the unintended consequence of assuming common knowledge of the payoff structure: coordination on multiple courses of action may not be possible when agents have different beliefs about the underlying fundamentals. In currency crises, for example, a unique equilibrium survives when speculators have heterogeneous information about the willingness and ability of the monetary authority to defend the currency.[2]

The comparative statics of the unique equilibrium suggest that policy instruments that affect agents' payoffs can be used to fashion market behavior. For example, raising domestic interest rates, restricting capital outflows, or borrowing reserves from abroad reduce the speculators' incentive to attack. Morris and Shin (1998) thus 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 about the policy maker's preferences, beliefs, and intentions. A central bank may be most anxious to raise interest rates when it is fearful of a large attack; conversely, not intervening may signal that the bank does not feel the need to take a preemptive strike.[3] Policy analysis therefore cannot be reduced to a simple comparative-static exercise – the issue is whether the market will interpret an intervention as a signal of resolve or a signal of distress.

This paper endogenizes policy in a global game of regime change that stylizes the role of coordination in applications.[4] A large number of agents choose whether to attack a status quo. A policy maker controls an instrument that affects the agents' payoff from attacking and maintains the status quo as long as the aggregate attack is small enough. In currency crises, regime change represents

[1] Examples include Diamond and Dybvig (1983), Obstfeld (1986, 1996), Katz and Shapiro (1986), Calvo (1988), Cooper and John (1988), Cole and Kehoe (2000), Battaglini and Benabou (2003). See Cooper (1998) for a review.

[2] See Morris and Shin (1998). Similar arguments have been made for bank runs (Goldstein and Pauzner, 2000; Rochet and Vives, 2004), debt crises (Corsetti, Guimaraes and Roubini, 2004), investment games (Chamley, 1999; Dasgupta, 2003), and riots (Atkeson, 2000).

[3] The idea that policy is a signal about the type of the policy maker goes back at least to Rogoff and Sibert (1988) and has been emphasized in the context of currency crises by Drazen (2000). This earlier work, however, does not examine coordination environments: the market is modeled as a single agent.

[4] Global games are games of incomplete information that often admit a unique equilibrium surviving iterated deletion of strictly dominated strategies; see Carlsson and van Damme (1993) for the pioneering contribution.

devaluation; in bank runs, the collapse of the banking system; in revolutions, the overturn of a dictator or some other sociopolitical establishment.

As in Morris and Shin, the policy maker's willingness or ability to maintain the status quo is not common knowledge, ensuring that the equilibrium would be unique with exogenous policy. Policy choices however depend on the type of the policy maker and can therefore convey valuable information to the market.

Our main result is that policy endogeneity leads to multiple equilibria. Different equilibria are sustained by the agents coordinating on different interpretations of, and different reactions, to the same policy choices.

There is an *inactive-policy equilibrium* where agents coordinate on a strategy that renders the aggregate attack insensitive to policy interventions, thus inducing the policy maker never to intervene. In addition, there is a continuum of *active-policy equilibria* where agents coordinate on the level of intervention at which they switch from aggressive to lenient behavior. The policy maker thus finds herself in a position where both the optimal policy and the regime outcome are dictated by self-fulfilling market expectations – a form of *policy trap* that contrasts sharply with Morris and Shin's policy prediction.

In the benchmark model, the policy is common knowledge and serves as a public signal about the type of the policy maker. Contrary to global games with exogenous public information (e.g., Morris and Shin, 2000; Hellwig, 2002), the informational content of this signal is endogenous and differs across equilibria. Nevertheless, one could argue that our multiplicity relies on the policy being publicly observed. This is not the case: the same multiplicity holds in perturbations of the game where the policy is observed with idiosyncratic noise.

Furthermore, multiplicity does not rely on the freedom to choose out-of-equilibrium beliefs. It is present even if agents have precise knowledge about the type of the policy maker and can be sustained in environments where beliefs are always pinned down by Bayes' rule.

Finally, our result is also different from the expectation traps of Chari, Christiano and Eichenbaum (1998) and Albanesi, Chari and Christiano (2003). In that work, multiplicity originates in the government's lack of commitment and vanishes if, as in our setting, the policy maker moves before the market.

The rest of the paper is organized as follows. Section 2 introduces the model. Section 3 presents the main results. Section 4 examines robustness to idiosyncratic noise in the observation of the policy. Section 5 discusses alternative payoff assumptions. Section 6 concludes. All formal proofs are in the Appendix.

2 The Model

Actions and payoffs. There are two possible regimes, a status quo and an alternative. A continuum of agents of measure one, indexed by i and uniformly distributed over $[0,1]$, choose whether to attack (i.e., take an action that favors regime change) or abstain from attacking (i.e. take an action that favors the status quo).

Let $r \in [\underline{r}, \bar{r}] \subset (0,1)$ denote the opportunity cost of attacking and $D \in \{0,1\}$ the regime outcome. The payoff for an agent who does not attack is normalized to zero, whereas the payoff from attacking is $1-r$ in the event the status quo is abandoned ($D=1$) and $-r$ otherwise ($D=0$).

Both r and D are controlled by a policy maker. As in Morris and Shin (1998), the payoff from maintaining the status quo is $V(\theta, A)$, where $\theta \in \mathbb{R}$ is the type of the policy maker (the "fundamentals"), $A \in [0,1]$ the mass of agents attacking, and V a continuous function, decreasing in A and increasing in θ, with $V(0,0) = V(1,1) = 0$. Hence, regime change is inevitable for $\theta < 0$; the status quo is sound but vulnerable to a sufficiently large attack for $\theta \in [0,1)$; and no attack can trigger regime change for $\theta \geq 1$. Finally, the cost of raising the policy to r is $C(r)$, where C is strictly increasing and Lipschitz continuous, with $C(\underline{r}) = 0$. The policy maker's payoff is therefore $U = (1-D)V(\theta, A) - C(r)$.

Information and timing. The game has three stages. In stage 1, the policy maker learns θ and sets r. In stage 2, agents decide simultaneously whether to attack after observing the policy r and receiving private signals about θ. Finally, in stage 3, the policy maker observes the aggregate attack A and decides whether to maintain the status quo.

The initial common prior about θ is an improper uniform over $\mathbb{R}$, whereas the signal that agent i receives is $x_i = \theta + \sigma\xi_i$.[5] $\sigma > 0$ parametrizes the quality of private information and ξ_i is idiosyncratic noise, i.i.d. across agents and independent of θ, with absolutely continuous c.d.f. Ψ and density ψ. We allow the support of the noise to be either $[-1,+1]$ (bounded) or the entire real line (unbounded) and denote with $\Theta(x) \equiv \{\theta : \psi(\frac{x-\theta}{\sigma}) > 0\}$ the set of types that are compatible with signal x. Bounded noise has the advantage that the freedom to choose out-of-equilibrium beliefs vanishes in the limit: $\Theta(x) \to \{x\}$ as $\sigma \to 0$ and hence an agent with signal x attaches probability one to $\theta = x$ no matter the strategy of the policy maker and the observed r. Unbounded noise, on the other hand, captures the possibility that agents may not be able to exclude any state.

To simplify the exposition, we let $V(\theta, A) = \theta - A$, in which case the policy maker finds it sequentially rational to maintain the status quo if and only if $A \leq \theta$ and hence her payoff reduces

[5] By assuming an uninformative prior, we bias the results *against* multiplicity and ensure that the equilibrium is unique with exogenous r for any $\sigma > 0$.

to $U(\theta, r, A) = \max\{0, \theta - A\} - C(r)$.[6] For future reference, we also define $\underline{x} = -\sigma$ and $\overline{x} = 1 + \sigma$ if ξ is bounded, $\underline{x} = -\infty$ and $\overline{x} = +\infty$ if ξ is unbounded, $\tilde{\theta} = 1 - \underline{r} \in (0, 1)$ and $\tilde{r} = C^{-1}(\tilde{\theta}) \in (\underline{r}, \overline{r}]$.[7]

Equilibrium. We consider symmetric perfect Bayesian equilibria. Let $r(\theta)$ denote the policy chosen by type θ, $\mu(\theta'|x, r)$ the agent's posterior belief that $\theta < \theta'$ given x and r, $a(x, r)$ the probability of attacking, and $A(\theta, r)$ the aggregate attack.

Definition. *An equilibrium consists of a policy function $r(\cdot)$, a strategy $a(\cdot)$ for the agents, and posterior beliefs $\mu(\cdot)$ such that:*

$$r(\theta) \in \arg\max_{r \in [\underline{r}, \overline{r}]} U(\theta, r, A(\theta, r)) \tag{1}$$

$$a(x, r) \in \arg\max_{a \in [0,1]} a\left[\int_{\Theta(x)} D(\theta, A(\theta, r))\, d\mu(\theta|x, r) - r\right] \tag{2}$$

$$\mu(\theta|x, r) \text{ is obtained from Bayes' rule using } r(\cdot) \text{ for any } r \in r(\Theta(x)) \tag{3}$$

where $A(\theta, r) = \int_{-\infty}^{+\infty} a(x, r)\psi\left(\frac{x-\theta}{\sigma}\right) dx$, $D(\theta, A) = 1$ *if* $A > \theta$, $D(\theta, A) = 0$ *if* $A \leq \theta$, *and* $r(\Theta(x)) \equiv \{r : r = r(\theta) \text{ for some } \theta \in \Theta(x)\}$.

Conditions (1) and (2) require that the policy choice in stage 1 and the agents' strategies in stage 2 are sequentially rational, while (3) requires that beliefs are pinned down by Bayes' rule along the equilibrium path. We also impose that out-of-equilibrium beliefs assign positive measure only to $\theta \in \Theta(x)$.

Remark. The defining elements of the model are the signaling game in stage 1 and the coordination game in stage 2; that stage 3 is strategic is not essential. As we discuss in Section 5, the results extend to more general environments where the regime outcome is not a choice of the policy maker.

3 Policy Traps

Suppose for a moment that the policy was exogenously fixed at some $r \in [\underline{r}, \overline{r}]$. The game then reduces to a standard global game with exogenous information structure, as in Morris and Shin (1998, 2003).

Proposition 1 *With exogenous policy, the equilibrium is unique: the status quo is abandoned if and only if $\theta < 1 - r$.*

[6] Without loss of generality, we assume that the policy maker maintains the status quo and that an agent attacks, when indifferent.

[7] Letting $\tilde{r} = C^{-1}(\tilde{\theta})$ presumes $C(\overline{r}) \geq \tilde{\theta}$; if the latter is not satisfied, the results hold with $\tilde{r} = \overline{r}$.

To prove this result, suppose that the status quo is abandoned if and only if $\theta < \theta^*$, for some $\theta^* \in (0,1)$. An agent with signal x then expects regime change with probability $\mu(\theta^*|x) = 1 - \Psi(\frac{x-\theta^*}{\sigma})$ and finds it optimal to attack if and only if $x < x^*$, where x^* solves $r = 1 - \Psi(\frac{x-\theta^*}{\sigma})$. It follows that the aggregate attack is $A(\theta, r) = \Psi\left(\frac{x^*-\theta}{\sigma}\right)$ and hence θ^* must solve $\theta^* = \Psi(\frac{x^*-\theta^*}{\sigma})$. Combining the two conditions gives $\theta^* = 1 - r$ and $x^* = 1 - r + \sigma\Psi^{-1}(1-r)$. Finally, iterated deletion of strictly dominated strategies selects this as the unique equilibrium of the game (see Appendix).

The comparative statics of Proposition 1 suggest that the policy maker can fashion market behavior and the regime outcome simply by undertaking policies that reduce the individual payoff from attacking: the higher is r, the lower x^* and θ^*.

This argument, however, fails to take into account that policy choices convey information. Consider for example the case of bounded noise. Since agents find it dominant to attack for $x < \underline{x}$, the size of attack is equal to one for $\theta < \underline{x} - \sigma$, no matter r, in which case it is dominant for the policy maker to set $\underline{r}$. Similarly, since agents find it dominant not to attack when $x > \overline{x}$, the policy maker necessarily sets $\underline{r}$ also for $\theta > \overline{x} + \sigma$. Any policy intervention thus signals that θ is neither too low nor too high – information that may interfere with the ability of the policy maker to control equilibrium outcomes.

Our main result is that, not only may interventions convey information, but agents can coordinate on multiple self-fulfilling expectations about the strategy of the policy maker and hence about the informational content of the same policy choices. The policy maker is thus trapped into a position where the best she can do is to conform to market expectations.

Theorem 1 *With endogenous policy, there are multiple equilibria: the level of policy intervention, the set of θ for which intervention occurs, and the range for which the status quo is abandoned, are indeterminate.*

We prove this result with Propositions 2 and 3 below.

3.1 Inactive-policy equilibrium

We first construct a pooling equilibrium in which agents coordinate on a strategy $a(x, r)$ that is insensitive to r, thus inducing the policy maker never to intervene.

Proposition 2 *There is an equilibrium in which the policy maker sets $\underline{r}$ for all θ, agents attack if and only if $x < \tilde{x}$, where $\tilde{x} = \tilde{\theta} + \sigma\Psi^{-1}(\tilde{\theta})$, and the status quo is abandoned if and only if $\theta < \tilde{\theta}$.*

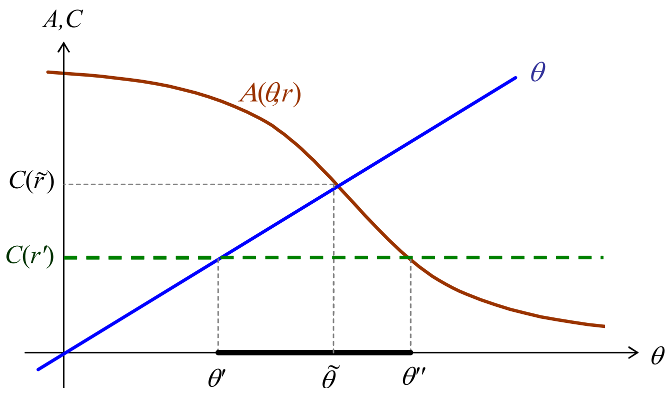

Figure 1: Inactive-policy equilibrium.

The construction of the inactive-policy equilibrium is illustrated in Figure 1. Given the agents' strategy, the status quo is abandoned if and only if $\theta < \tilde{\theta}$, no matter r, where $\tilde{\theta}$ solves $\tilde{\theta} = A(\tilde{\theta}, r)$. When $r = \underline{r}$, beliefs are pinned down by Bayes rule and satisfy $\mu(\tilde{\theta}|x, \underline{r}) > \underline{r}$ if and only if $x < \tilde{x}$. For $r \neq \underline{r}$, consider out-of-equilibrium beliefs that assign zero measure to types for whom the deviation is dominated in equilibrium (whenever possible). Since the policy maker's equilibrium payoff is 0 for $\theta \leq \tilde{\theta}$ and $\theta - A(\theta, \underline{r}) > 0$ for $\theta > \tilde{\theta}$, any $r > \tilde{r}$ is dominated in equilibrium for all θ, in which case the only restriction is that beliefs have support $\Theta(x)$. A deviation to some $r' \in (\underline{r}, \tilde{r})$, on other hand, is dominated if and only if $\theta \notin [\theta', \theta'']$, where θ' and θ'' solve $\theta' = C(r') = A(\theta'', \underline{r})$: for $\theta < \theta'$ the cost of r' exceeds the value of maintaining the status quo, whereas for $\theta > \theta''$ the attack faced in equilibrium is smaller than the cost of r'. Since $\tilde{\theta} \in [\theta', \theta'']$, beliefs may assign zero measure to $\theta \notin [\theta', \theta'']$ whenever $[\theta', \theta''] \cap \Theta(x) \neq \varnothing$ and at the same time satisfy $\mu(\tilde{\theta}|x, r) > r$ if and only if $x < \tilde{x}$, for all r. Given these beliefs, an agent who expects all other agents to follow the proposed strategy, thus triggering regime change if and only if $\theta < \tilde{\theta}$, finds its optimal to do the same. The size of attack $A(\theta, r)$ is then independent of r, eliminating any incentive for policy intervention.

Clearly, any beliefs and strategies such that $A(\theta, r) \geq A(\theta, \underline{r})$ for any (θ, r) sustain policy inaction as an equilibrium. The ones considered here have two advantages. First, the beliefs satisfy a simple forward-induction argument as in Cho and Kreps' (1987) intuitive criterion. Second, the associated strategies are the limit of those that implement policy inaction when r is observed with unbounded idiosyncratic noise, in which case there is no room for out-of-equilibrium beliefs (see Section 4.2).

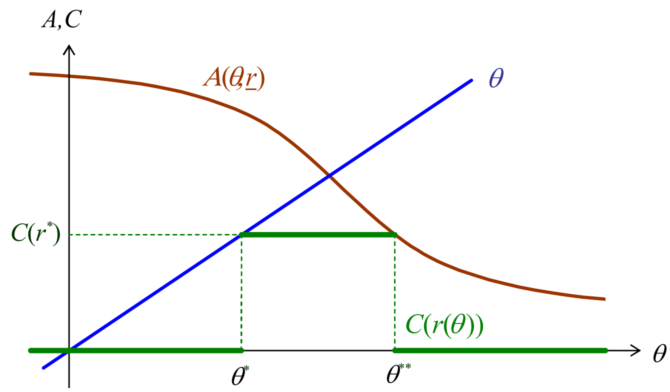

Figure 2: Active-policy equilibrium.

3.2 Active-policy equilibria

Suppose now agents coordinate on a strategy $a(x,r)$ that is decreasing not only in x but also in r. In this case, raising r can decrease the size of attack and even preempt regime change. However, the cost of intervention may exceed the value of maintaining the status quo for low θ; similarly, the attack faced when setting $\underline{r}$ may be too small to justify intervention when θ is high. This suggests the existence of equilibria in which intervention occurs only for intermediate θ.

Proposition 3 *For any $r^* \in (\underline{r}, \tilde{r}]$, there is an equilibrium in which the policy maker sets $r(\theta) = r^*$ if $\theta \in [\theta^*, \theta^{**}]$ and $r(\theta) = \underline{r}$ otherwise, agents attack if and only if $x < \underline{x}$ or $(x,r) < (x^*, r^*)$, and the status quo is abandoned if and only if $\theta < \theta^*$, where*

$$\theta^* = C(r^*), \quad \theta^{**} = \theta^* + \sigma[\Psi^{-1}(1 - \tfrac{\underline{r}}{1-\underline{r}}\theta^*) - \Psi^{-1}(\theta^*)], \quad x^* = \theta^{**} + \sigma\Psi^{-1}(\theta^*). \tag{4}$$

The construction of an active-policy equilibrium is illustrated in Figure 2. When the agents coordinate on the equilibrium strategy, it never pays to raise the policy at any $r \neq r^*$. Furthermore, r^* is preferred to $\underline{r}$ if and only if $C(r^*) \leq \theta$ and $C(r^*) \leq A(\theta, \underline{r})$. The thresholds θ^* and θ^{**} thus solve $\theta^* = C(r^*) = \Psi\left(\frac{x^* - \theta^{**}}{\sigma}\right)$, while x^* solves the indifference condition for the agents, $\underline{r} = \mu(\theta^* | x^*, \underline{r})$. Combining these three conditions gives (4), while ensuring that the set of types who intervene is non-empty puts an upper bound on the cost of intervention and hence on r^*.

Other strategies can sustain the same policy. For example, agents could coordinate on attacking if and only if $x < \overline{x}$ whenever the policy maker does not conform to market expectations, that is, whenever the policy is raised at some $r \neq r^*$. As with Proposition 2, the proposed strategies are sustained by beliefs that assign zero measure to types for whom a deviation is dominated in equilibrium and are the limit of those in a perturbed game where beliefs are always pinned down by Bayes' rule (see Section 4.2).

Also note that the exact θ is never revealed in any of the above equilibria; the equilibrium policy only makes it common certainty whether $\theta \in [\theta^*, \theta^{**}]$ or $\theta \notin [\theta^*, \theta^{**}]$.[8] This type of common certainty permits perfect coordination on not attacking when the policy is raised at r^*. As discussed in the next section, however, such a form of perfect coordination is not essential for multiplicity.

Finally, contrast Proposition 2 with the case where a single "big" agent plays against the policy maker. This agent recognizes that he can trigger regime change for all $\theta < 1$. When noise is bounded, he necessarily attacks for $x < 1 - \sigma$. As $\sigma \to 0$, the status quo is thus abandoned if and only if $\theta < 1$. In our coordination setting, instead, the regime outcome remains indeterminate for any $\theta \in (0, \tilde{\theta}]$.

4 Idiosyncratic policy observation

The payoff structure of the coordination game played among the agents depends on two variables, θ and r. We have assumed that, while θ is observed with idiosyncratic noise, r is observed publicly. Although this is a reasonable assumption for most applications, from a global-games perspective it is important to consider perturbations that remove common knowledge of the policy.[9]

In this section, we consider two such perturbations. In the first, the policy is observed with small bounded idiosyncratic noise; in the second, the support of the policy signals does not shift with the actual policy choice. In both cases, there is no public information about either r or θ.

The key difference from standard global games is the endogeneity of the information structure: whereas the informational content of the private signals about θ is exogenously given, that of the signals about r is determined in equilibrium. As we show next, this has important implications for the determinacy of equilibria.

4.1 Bounded policy noise

Consider the following modification of the benchmark model. In stage 2, each agent receives a private signal $z_i = r + \eta\zeta_i$ about the policy; $\eta > 0$ parametrizes the precision of the policy signal and ζ_i is noise, i.i.d. across agents and independent of θ and ξ_i, distributed over $[-1, 1]$ with absolutely continuous c.d.f. Φ and density ϕ bounded away from 0.

[8] By "common certainty" we mean common $p = 1$ beliefs (Monderer and Samet, 1989).

[9] Another possibility is that r is observed with aggregate noise. This case may be relevant for some applications, but is less interesting from a theoretical perspective, for it maintains common knowledge of the policy signal.

Proposition 4 *Multiplicity survives with small bounded policy noise.*

(i) There is an inactive-policy equilibrium as in Proposition 2.

(ii) There exists $r_\eta > \underline{r}$*, with* $r_\eta \to \underline{r}$ *as* $\eta \to 0$*, such that, for any* $r^* \in (r_\eta, \tilde{r}]$*, there is an active-policy equilibrium as in Proposition 3.*

In the inactive-policy equilibrium, an agent who observes a signal $z \in [\underline{r} - \eta, \underline{r} + \eta]$ believes that the policy was set at $\underline{r}$ and that all other agents also received signals $z \in [\underline{r} - \eta, \underline{r} + \eta]$. Attacking if and only if $x < \tilde{x}$ is then sequentially rational for an agent who expects all other agents to do the same. Moreover, the same strategy can be sustained for $z > \underline{r} + \eta$ by beliefs that satisfy the intuitive criterion.[10]

In an active-policy equilibrium, on the other hand, agents coordinate on attacking if and only if $x < \underline{x}$ or $(x, z) < (x^*, z^*)$, where x^* is as in Proposition 3 and $z^* = r^* - \eta > \underline{r} + \eta$ (for η sufficiently small). Since the size of attack is the same for all $r < z^* - \eta$, the policy marker never sets $r \in (\underline{r}, r^* - 2\eta)$. Moreover, while the marginal benefit of reducing r below r^* is independent of η, the marginal increase in the size of attack goes to infinity as η goes to zero. Hence, for η small enough, the policy maker sets either $\underline{r}$ or r^*. Similar arguments as in Proposition 3 then imply that r^* is optimal if and only if $\theta \in [\theta^*, \theta^{**}]$.

Although r is not publicly observed, in an active-policy equilibrium the observation of $z \in [\underline{r} - \eta, \underline{r} + \eta] \cup [r^* - \eta, r^* + \eta]$ generates common certainty on whether $r = \underline{r}$ or $r = r^*$, and hence on whether θ is extreme or intermediate. This however is not a consequence of bounded noise alone: if the equilibrium policy had no discontinuities, no policy choice would ever lead to common certainty about either r or θ.

4.2 Unbounded policy noise

One may argue that multiplicity survives with small bounded noise only because of the common certainty generated by equilibrium policies. Moreover, agents can still detect deviations. As in standard signaling games, one may thus argue that multiplicity relies on the freedom to choose out-of-equilibrium beliefs.

To show that neither of the above is necessarily true, we consider the following example. The private signal about the policy is $z = w\underline{r} + (1 - w)r + \eta\zeta$, where w is a binary variable assuming

[10] Note that this is a game with noisy signaling. Forward induction puts restrictions not only on the beliefs about θ but also about r : if $r \in [z - \eta, z + \eta]$ is dominated in equilibrium for all $\theta \in \Theta(x)$, whereas there is $r' \in [z - \eta, z + \eta]$ that is not dominated for some $\theta \in \Theta(x)$, then the prescribed strategy profile should not rely on the agents assigning positive measure to r.

value 1 with probability $\rho \in (0,1)$ and 0 otherwise, whereas ζ is distributed exponentially over $[0,+\infty)$ and $\eta > 0$. The noises w and ζ are i.i.d. across agents and independent of θ, ξ, and each other. That ζ is exponential simplifies the construction of equilibria by ensuring that, when the policy takes only two values, their likelihood ratio conditional on z also takes only two values; w then ensures that the support of the policy signal is $\mathcal{Z} = [\underline{r}, \infty)$ for any policy choice.[11] We also assume that C is linear and ψ is log-concave and strictly positive over $\mathbb{R}$. The combination of these assumptions keeps the analysis tractable.

Proposition 5 *Consider the noise structure described above.*

(i) There is an inactive-policy equilibrium as in Proposition 2.

(ii) Any of the active-policy equilibria in Proposition 3 can be approximated by an equilibrium in the perturbed game: for any $r^* \in (\underline{r}, \tilde{r})$, $\varepsilon > 0$, and (η, ρ) small enough, there is an equilibrium in which $r(\theta) = r^*$ if $\theta \in [\theta', \theta'']$, $r(\theta) = \underline{r}$ otherwise, and the status quo is abandoned if and only if $\theta < \theta'$, with $|\theta' - \theta^*| < \varepsilon$, $|\theta'' - \theta^{**}| < \varepsilon$, and (θ^*, θ^{**}) as in (4).

Since any (x,z) is consistent with any (θ, r), beliefs are always pinned down by Bayes' rule and no policy ever generates certainty – either private or common – about the fundamentals. Hence, in contrast to both the benchmark model and the bounded-noise case, agents can no longer perfectly coordinate on not attacking whenever the policy is set sufficiently high. Indeed, an agent necessarily attacks when x is low enough, no matter z.

Nevertheless, agents can still coordinate on different interpretations of, and different reactions to, the same idiosyncratic policy signals.

In an inactive-policy equilibrium, agents expect the policy maker never to intervene and interpret variation in z as pure noise. They then condition their behavior only on x, thus making the aggregate attack independent of r.

In an active-policy equilibrium, instead, an agent who observes a high z attaches high probability to other agents also having observed high policy signals. Hence, if he expects other agents to play more leniently when they observe sufficiently high z, he finds it optimal to do the same. But how high z needs to be for an agent to play more leniently – and therefore the optimal level of policy intervention – depends again on market expectations.

In other words, whereas in the baseline model agents coordinate on the sensitiveness of their strategies to the public signal r, now they coordinate on the sensitivity to the private signals z.

[11] The same multiplicity can be sustained when $\rho = 0$.

Remark. The two noise structures considered above allow to sustain exactly the same type of equilibria as in the benchmark model. Although this may not be possible with other noise structures, the logic of multiplicity does not seem to be specific to these examples.

5 Discussion

The benchmark model assumes that the cost of policy intervention is independent of θ; it also identifies the strength of the status quo with the value the policy maker attaches to it. These assumptions, however, might not be appropriate for some applications. Similarly, some policies may improve upon the strength of the status quo instead of raising the agents' cost of attacking, as in the case of a central bank borrowing reserves from abroad.

To capture these possibilities, we extend the game as follows. The status quo is maintained ($D = 0$) if and only if $R(\theta, A, r) \geq 0$, where R is continuous in (θ, A, r), strictly increasing in θ, strictly decreasing in A, and non-decreasing in r, with $R(0, 0, r) = 0 = R(1, 1, r)$ for any r. When $D = 0$, the policy maker's payoff is given by $U(\theta, A, r)$, where U is continuous in (θ, A, r), non-decreasing in θ, non-increasing in A, and strictly decreasing in r. When instead $D = 1$, her payoff is 0 if $r = \underline{r}$ and $W(\theta, A, r) < 0$ if $r > \underline{r}$.[12] Finally, the agent's net payoff from attacking is $1 - g(r)$ if $D = 1$ and $-g(r)$ otherwise, where g is non-decreasing in r, with $0 < g(\underline{r}) \leq g(\overline{r}) < 1$. The benchmark model is nested with $R(\theta, A, r) = V(\theta, A)$, $U(\theta, A, r) = V(\theta, A) - C(r)$, and $g(r) = r$.

Let $\tilde{\theta} \in (0, 1)$ and $\tilde{r} \in (\underline{r}, \overline{r})$ solve $R(\tilde{\theta}, 1 - g(\underline{r}), \underline{r}) = 0$ and $U(\tilde{\theta}, 0, \tilde{r}) = 0$.[13] Assume finally that ψ is log-concave. The existence of an inactive-policy equilibrium as in Proposition 2 is straightforward; the following generalizes Proposition 3.

Proposition 6 *Consider the extension described above. For any $r^* \in (\underline{r}, \tilde{r})$, there is a non-empty set Θ^{**} and an equilibrium in which the policy is $r(\theta) = r^*$ if $\theta \in \Theta^{**}$ and $r(\theta) = \underline{r}$ otherwise and the status quo is abandoned if and only if $\theta < \theta^*$, where $\theta^* = \min \Theta^{**} \in [0, \tilde{\theta})$.*

Unlike the equilibria of Proposition 3, the set of fundamentals for which the policy maker intervenes may now be the union of multiple disjoint intervals. For example, let $R(\theta, A, r) = \theta - A$ and $U(\theta, A, r) = \theta - A - C(r, \theta)$, where $C(r, \theta)$ is decreasing in θ for $r > \underline{r}$. The equation $C(r, \theta) = A(\theta, \underline{r})$ may then admit multiple solutions corresponding to multiple indifference points at which the policy maker switches between r^* and $\underline{r}$ without facing regime change.

[12] Note that r does not affect the boundaries of the critical region and $r = \underline{r}$ is optimal when the status quo is abandoned. These assumptions simplify the proof of Proposition 6. We do not expect the multiplicity result to be unduly sensitive to these restrictions.

[13] This presumes $U(\tilde{\theta}, 0, \underline{r}) > 0 > U(\tilde{\theta}, 0, \overline{r})$, i.e., that $\tilde{\theta}$ would like to escape regime change by raising r.

Moreover, there may exist equilibria in which the policy is raised to r^* for all $\theta \geq \theta^*$. Such equilibria are sustained by the agents coordinating on attacking if and only if $r < r^*$ no matter x, and therefore require that the noise ξ is unbounded.[14] Alternatively, it may be the lowest types in the critical region who raise the policy. For example, if $R(\theta, A, r) = \theta - A$ and $U(\theta, A, r) = V - C(r)$, where $V > 0$ is a constant, then $\Theta^{**} = [0, \theta^{**})$ for some $\theta^{**} > 0$.

A possibility not addressed by the above result is that the strength and the value of the status quo are negatively correlated with each other – countries with the weakest fundamentals might be those that suffer the most from a collapse of the currency or the banking system. We consider an example in the online appendix (Section A1). We find again multiple equilibria in which policy intervention occurs for the lowest types in the critical region.

These results suggest that multiplicity is likely to extend to a variety of applications. At the same time, the global-games methodology does not necessarily loose all its selection power.

This is most evident in the following example. As in the benchmark model, let $R(\theta, A, r) = V(\theta, A)$, $U(\theta, A, r) = V(\theta, A) - C(r)$, and $g(r) = r$; but now assume $V_{\theta A} \geq 0$ and $\lim_{\theta \to \infty}[V(\theta, 0) - V(\theta, 1)] = 0$, meaning that the cost of an attack is non-increasing in θ and vanishes as $\theta \to \infty$. Suppose further that the noise ξ is unbounded and has log-concave density. These assumptions ensure that sufficiently high θ do not intervene, that there is at most one policy level other than $\underline{r}$ played in equilibrium, and that the agents' strategy is monotonic in x. We can then show that the entire set of equilibrium outcomes is given by the equilibria of Propositions 2 and 3 (see Section A2 in the online appendix).

The following predictions can thus be made irrespectively of the equilibrium played: first, the regime outcome is monotonic in θ and the status quo is necessarily maintained for $\theta > \tilde{\theta}$; second, policy interventions take place only for an intermediate region of types; and third, this region vanishes as agents become perfectly informed (i.e., as $\sigma \to 0$). In contrast, none of these predictions are possible when θ is common knowledge: policy intervention and regime change can then occur for any subset of the critical range.

We conclude that incomplete information may significantly reduce the set of equilibrium outcomes and possibly lead to interesting predictions despite multiplicity. These predictions however are sensitive to the details of the payoff structure and therefore can be appreciated only within the context of specific applications.

[14] When the noise is bounded, it is dominant for the agent not to attack whenever $x > \bar{x}$ and hence for the policy maker to set $\underline{r}$ whenever $\theta > \bar{x} + \sigma$. See the discussion of "one-threshold equilibria" in the working-paper version of this article (Angeletos, Hellwig and Pavan, 2003).

Finally, for some applications of interest, θ may represent some underlying economic fundamentals about which the policy maker has only imperfect information. In the online appendix (Section A3) we show that multiplicity is robust to small bounded noise in the policy maker's observation of θ. The case of large noise is perhaps more interesting, for it may introduce novel effects such as strategic experimentation and learning. These issues however are beyond the scope of this paper and are left for future research.

6 Concluding remarks

This paper endogenized policy in a global coordination game. We found that the possibility that policy choices convey information leads to multiple equilibria where both the optimal policy and the coordination outcome are dictated by self-fulfilling market expectations. Different equilibria are sustained by the agents coordinating on multiple interpretations of and reactions to the same policy choices.

On the theoretical side, our results underscore the importance of endogenous information structures in global games. On the applied, they raise questions about the merits of certain policy proposals.

Can a central bank prevent a financial crisis by injecting liquidity, or will that be interpreted as a signal of distress? And do IMF interventions ease debt crises, or might they do more harm than good by revealing that country fundamentals are weak enough to require IMF aid? To address these questions, one has to examine the informational role of policy.

Appendix

Proof of Proposition 1 Starting from $\underline{\theta}_0 \equiv 0$, construct a sequence $\{\underline{x}_k, \underline{\theta}_k\}_{k=0}^{\infty}$ by letting $\underline{x}_k$ and $\underline{\theta}_{k+1}$ be the unique solutions to $r = 1 - \Psi(\frac{\underline{x}_k - \underline{\theta}_k}{\sigma})$ and $\underline{\theta}_{k+1} = \Psi(\frac{\underline{x}_k - \underline{\theta}_{k+1}}{\sigma})$. This sequence represents iterated dominance from below: whatever the strategy of other agents, the posterior probability of regime change is at least $1 - \Psi(\frac{x-0}{\sigma})$ and hence it is dominant to attack for $x < \underline{x}_0$; conditional on others attacking when $x < \underline{x}_0$, the probability of regime change is then at least $1 - \Psi(\frac{x - \underline{\theta}_1}{\sigma})$ making it dominant to attack for $x < \underline{x}_1$; and so on. Moreover, this sequence is increasing and bounded from above, which together with the continuity of Ψ implies that $\lim_{k\to\infty}(\underline{x}_k, \underline{\theta}_k) = (x^*, \theta^*)$, where

$\theta^* = 1 - r$ and $x^* = 1 - r + \sigma\Psi^{-1}(1-r)$ are the thresholds corresponding to the unique monotone equilibrium. It follows that necessarily $a(x,r) = 1$ for all $x < x^*$ and $D(\theta) = 1$ for all $\theta < \theta^*$. A symmetric argument from above establishes that $a(x,r) = 0$ for all $x > x^*$ and $D(\theta) = 0$ for all $\theta > \theta^*$. *QED*

Proof of Proposition 2. When $r = \underline{r}$, beliefs are pinned down by Bayes rule and, since the observation of $\underline{r}$ is uninformative about θ, they are given by $\mu(\theta|x,\underline{r}) = 1 - \Psi(\frac{x-\theta}{\sigma})$. Note that, by definition of $\tilde{\theta}$ and $\tilde{x}$, $\mu(\tilde{\theta}|x,\underline{r}) > \underline{r}$ if and only if $x < \tilde{x}$. When instead $r > \underline{r}$, consider out-of-equilibrium beliefs as follows. Let $\Theta(r)$ be the set of types for whom r is dominated in equilibrium: $\Theta(r) \equiv \{\theta : U^*(\theta) > U(\theta, r, \Psi(\frac{x-\theta}{\sigma}))\}$, where $U^*(\theta) \equiv U(\theta, \underline{r}, \Psi(\frac{\tilde{x}-\theta}{\sigma}))$. For $r > \tilde{r}$, $\Theta(r) = \mathbb{R}$; for $r = \tilde{r}$, $\Theta(r) = \mathbb{R}\backslash\{\tilde{\theta}\}$; and for $r \in (\underline{r}, \tilde{r})$, $\Theta(r) = (-\infty, \theta') \cup (\theta'', +\infty)$, where θ' and θ'' solve $\theta' = C(r) = \Psi(\frac{\tilde{x}-\theta''}{\sigma})$ and satisfy $0 < \theta' < \tilde{\theta} < \theta'' < \infty$. Then restrict μ so that $\mu(\{\theta \in \Theta(x) \cap \Theta(r)\}|x,r) = 0$ whenever $\Theta(x) \nsubseteq \Theta(r)$.[15] If $r = \tilde{r}$ and $\tilde{\theta} \in \Theta(x)$, in which case $\mu(\{\theta = \tilde{\theta}\}|x,r) = 1$, assume that type $\tilde{\theta}$ – who is indifferent between maintaining and abandoning – is expected to abandon with probability $\tilde{r}$. In all other cases, let $\mu(\tilde{\theta}|x,r) > r$ if and only if $x < \tilde{x}$. Given these beliefs, an agent who expects all other agents to follow the proposed strategy finds it optimal to do the same. Since C is strictly increasing and the equilibrium $A(\theta, r)$ does not depend on r, any θ then clearly finds it optimal to set $\underline{r}$. *QED*

Proof of Proposition 3. Take any $r^* \in (\underline{r}, \tilde{r}]$ and note that $0 < \theta^* \leq \tilde{\theta}$. When agents coordinate on the proposed strategy, the status quo is abandoned if and only if $\theta < 0$ or $(r,\theta) < (r^*, \hat{\theta})$, where $\hat{\theta}$ solves $\hat{\theta} = \Psi(\frac{x^*-\hat{\theta}}{\sigma})$ and $\theta^* \leq \hat{\theta} \leq \theta^{**} < -\infty$, with the equalities holding only for $r^* = \tilde{r}$.

First, consider the behavior of an agent. When $r = \underline{r}$, beliefs are pinned down by Bayes' rule, since (4) ensures $\Theta(x) \nsubseteq [\theta^*, \theta^{**}]$ for all x. (When the noise is unbounded, this is immediate; when it is bounded, it follows from $|\theta^{**} - \theta^*| < 2\sigma$.) The posterior belief of regime change is

$$\mu(\hat{\theta}|x,\underline{r}) = \mu(\theta^*|x,\underline{r}) = \frac{1 - \Psi(\frac{x-\theta^*}{\sigma})}{1 - \Psi(\frac{x-\theta^*}{\sigma}) + \Psi(\frac{x-\theta^{**}}{\sigma})}$$

and, by (4), $\mu(\theta^*|x^*,\underline{r}) > \underline{r}$ if and only if $x < x^*$. When $r \in (\underline{r}, r^*)$, the set of types for whom r is dominated in equilibrium is $\Theta(r) = (-\infty, \theta') \cup (\theta'', +\infty)$, where θ' and θ'' solve $\theta' = C(r) = \Psi\left(\frac{x^*-\theta}{\sigma}\right)$ and satisfy $0 < \theta' < \theta^* \leq \hat{\theta} \leq \theta^{**} < \theta'' < \infty$. Then, take any beliefs μ such that $\mu(\hat{\theta}|x,r) > r$ if and only if $x < x^*$ and $\mu(\{\theta \in \Theta(x) \cap \Theta(r)\}|x,r) = 0$ whenever $\Theta(x) \nsubseteq \Theta(r)$. When instead $r = r^*$ and $\Theta(x) \cap [\theta^*, \theta^{**}] \neq \varnothing$, Bayes' rule implies $\mu(0|x,r) = 0$. Finally, for any (x,r) such that either

[15] With some abuse of notation, $\mu(\{E\}|x,r)$ denotes the posterior probabiltiy of event E.

$r > r^*$ or $r = r^*$ and $\Theta(x) \cap [\theta^*, \theta^{**}] = \varnothing$, in which case $\Theta(r) = \Theta(x)$, take any beliefs such that $\mu(0|x, r) = 0$ if $x \geq \underline{x}$ and $\mu(0|x, r) = 1$ otherwise. Given these beliefs, the proposed strategy is sequentially rational.

Next, consider the policy maker. Given the agents' strategy, $\underline{r}$ is preferred to any $r \in (\underline{r}, r^*)$ and r^* to any $r > r^*$. For $\theta < \theta^*$, the payoff from setting r^* is $\theta - C(r^*) < 0$ and hence $\underline{r}$ is optimal. For $\theta \in (\theta^*, \hat{\theta})$, r^* is optimal, since $\underline{r}$ leads to regime change whereas r^* yields $\theta - C(r^*) > 0$. For $\theta > \hat{\theta}$, the status quo survives even if $r = \underline{r}$, but since $A(\theta, \underline{r}) > C(r^*)$ if and only if $\theta < \theta^{**}$, it is optimal to set r^* for $\theta \in (\hat{\theta}, \theta^{**})$ and $\underline{r}$ for $\theta > \theta^{**}$.

Finally, note that $\theta^* \leq \theta^{**}$ if and only if $\theta^* \leq 1 - \underline{r}$ $(= \tilde{\theta})$ and therefore an active-policy equilibrium of the type considered above exists if and only if $\theta^* \in (0, \tilde{\theta}]$, or equivalently $r^* \in (\underline{r}, \tilde{r}]$, which completes the proof. *QED*

Proof of Proposition 4. Since C is Lipschitz continuous, there is $K < \infty$ such that $|C(r) - C(r')| < K|r - r'|$ for any $(r, r') \in \mathcal{R}^2$, where $\mathcal{R} \equiv [\underline{r}, \overline{r}]$. We prove the result for $\eta < \bar{\eta}$, where $\bar{\eta} = \min\{(\tilde{r} - \underline{r})/4; (1 - \underline{r})\underline{\phi}/K)\}$ and $\underline{\phi} = \inf_{\zeta \in [-1,1]} \phi(\zeta) > 0$. With a slight abuse of notation, we denote with $\mu(\{r = r'\}|x, z)$ the posterior probability that an agent with information (x, z) assigns to the event that $r = r'$ and with $\mu(\{\theta \in \Theta', r = r'\}|x, z)$ the joint probability that $\theta \in \Theta'$ and $r = r'$.

Part (i). We prove that there exists an equilibrium in which $r(\theta) = \underline{r}$ for all θ, agents attack if and only if $x < \tilde{x}$, whatever z, and the status quo is abandoned if and only if $\theta < \tilde{\theta}$.

Consider first the agents. When $z \leq \underline{r} + \eta$, beliefs are necessarily pinned down by Bayes' rule and sequential rationality for the agents follows directly from the same arguments as in the proof of Proposition 2.

When instead $z > \underline{r} + \eta$, the prescribed strategy can be sustained with out-of-equilibrium beliefs that guarantee that the equilibrium passes the intuitive criterion test. Let $\Theta(r')$ denote the set of types for whom r' is dominated in equilibrium. For $r' > \tilde{r}$, $\Theta(r') = \mathbb{R}$; for $r = \tilde{r}$, $\Theta(r') = \mathbb{R}/\{\tilde{\theta}\}$; and for $r' \in (\underline{r} + 2\eta, \tilde{r})$, $\Theta(r') = \left(-\infty, \theta'\right) \cup \left(\theta'', +\infty\right)$, where θ' and θ'' solve $\theta' = C(r) = A(\theta'', \underline{r})$, with $A(\theta, \underline{r}) = \Psi(\frac{\tilde{x} - \theta}{\sigma})$.[16] For any (x, z), then let $P(x, z) = \{r' \in [z - \eta, z + \eta] \cap \mathcal{R} : \Theta(x) \nsubseteq \Theta(r')\}$ denote the set of policies that are compatible with z and which are *not* dominated in equilibrium for some $\theta \in \Theta(x)$.

Take first any (x, z) such that $\tilde{\theta} \notin \Theta(x)$, which is possible only when ξ is bounded. If $P(x, z) \neq$

[16] Note that for all $\theta \in [\theta', \theta'']$, $x > \underline{x}$. Hence, an agent with signals (x, z) with $z > \underline{r} + \eta$ who believes that $r' > \underline{r} + 2\eta$ and who expects all other agents not to attack, also finds it optimal not to attack. It follows that the minimum size of attack for a type $\theta \in [\theta', \theta'']$ who sets the policy at $r' \in (\underline{r} + 2\eta, \tilde{r})$ is $A(\theta, r') = 0$, which implies that r' is not dominated in equilibrium.

$\varnothing$, then pick any $\rho \in P(x,z)$ and let μ be any belief that satisfies $\mu(\{r = \rho\}|x,z) = 1$ and $\mu(\{\theta \in \Theta(x) \cap \Theta(\rho), r = \rho\}|x,z) = 0$. If instead $P(x,z) = \varnothing$, then simply take any $\rho \in [z - \eta, z+\eta] \cap \mathcal{R}$ and any μ such that $\mu(\{r = \rho\}|x,z) = 1$. Note that $\tilde{\theta} \notin \Theta(x)$ implies that either $x < \tilde{\theta} - \sigma < \tilde{x}$ in which case μ also satisfies $\mu(\{\theta < \tilde{\theta}, r = \rho\}|x,z) = 1$, or $x > \tilde{\theta} + \sigma > \tilde{x}$ in which case $\mu(\{\theta < \tilde{\theta}, r = \rho\}|x,z) = 0$.

Take next any (x,z) such that $\tilde{\theta} \in \Theta(x)$. If $z \in (\underline{r} + \eta, \tilde{r} + \eta)$, then necessarily $P(x,z) \neq \varnothing$ since any $\rho \in [z-\eta, z+\eta] \cap (\underline{r} + 2\eta, \tilde{r})$ is never dominated in equilibrium for $\theta \in [\theta'(\rho), \theta''(\rho)]$, where $\theta'(\rho)$ and $\theta''(\rho)$ solve $\theta' = C(\rho) = \Psi(\frac{\tilde{x} - \theta''}{\sigma})$ and satisfy $0 < \theta'(\rho) < \tilde{\theta} < \theta''(\rho) < \infty$. Then, take any $\rho \in [z-\eta, z+\eta] \cap (\underline{r} + 2\eta, \tilde{r})$ and let μ be any beliefs such that $\mu(\{r = \rho\}|x,z) = 1$, $\mu(\{\theta \in \Theta(x) \cap \Theta(\rho), r = \rho\}|x,z) = 0$ and $\mu(\{\theta < \tilde{\theta}, r = \rho\}|x,z) > \rho$ if and only if $x < \tilde{x}$. If instead $z > \tilde{r} + \eta$, necessarily $P(x,z) = \varnothing$. Then simply take any $\rho \in [z - \eta, z+\eta] \cap \mathcal{R}$ and any μ such that $\mu(\{r = \rho\}|x,z) = 1$ and $\mu(\{\theta < \tilde{\theta}, r = \rho\}|x,z) > \rho$ if and only if $x < \tilde{x}$. Finally, if $z = \tilde{r} + \eta$, since $P(x,z) = \{\tilde{r}\}$ and $\Theta(x) \cap \Theta(\tilde{r}) = \Theta(x)/\{\tilde{\theta}\}$, then let $\mu(\{\theta = \tilde{\theta}, r = \tilde{r}\}|x,z) = 1$, and, as in the proof of Proposition 2, assume this agent also expects $\tilde{\theta}$ to abandon the regime with probability $\tilde{r}$.

With the beliefs specified above, an agent who expects all other agents to attack if and only $x < \tilde{x}$ for any z, finds it optimal to follow the same strategy.

Finally, since the size of attack does not depend on r, setting $r(\theta) = \underline{r}$ for all θ is sequentially rational for the policy maker.

Part (ii). Let $r_\eta \equiv \max\{\underline{r} + 4\eta,\ C^{-1}(K\eta/\underline{\phi})\}$ and note that r_η is increasing in η, $r_\eta \to \underline{r}$ as $\eta \to 0$, and $r_\eta < \tilde{r}$ for any $\eta < \bar{\eta}$. Take any $r^* \in (r_\eta, \tilde{r}]$ and let $z^* = r^* - \eta$ and $(x^*, \theta^*, \theta^{**})$ as in (4). We prove that there exists an equilibrium in which $r(\theta) = r^*$ if $\theta \in [\theta^*, \theta^{**}]$ and $r(\theta) = \underline{r}$ otherwise, agents attack if and only if $x < \underline{x}$ or $(x,r) < (x^*, z^*)$, and the status quo is abandoned if and only if $\theta < \theta^*$.

Consider first the agents. When $z \in [\underline{r} - \eta, \underline{r} + \eta]$, beliefs are always pinned down by Bayes' rule, since (4) ensures $\Theta(x) \nsubseteq [\theta^*, \theta^{**}]$ for any x. It follows that $\mu(\{r = \underline{r}\}|x,z) = 1$ and

$$\mu(\{\theta < \hat{\theta}, r = \underline{r}\}|x,z) = \mu(\{\theta < \theta^*, r = \underline{r}\}|x,z) = \frac{1 - \Psi(\frac{x - \theta^*}{\sigma})}{1 - \Psi(\frac{x-\theta^*}{\sigma}) + \Psi(\frac{x - \theta^{**}}{\sigma})},$$

where $\hat{\theta} \in (\theta^*, \theta^{**})$ solves $\Psi(\frac{x^* - \hat{\theta}}{\sigma}) = \hat{\theta}$ and x^* is as in (4).

Take any $z \in (\underline{r} + \eta, r^* - \eta)$. If $\Theta(x) \cap [\theta^*, \theta^{**}] \neq \varnothing$, pick any $\rho \in (\underline{r} + 2\eta, r^* - 2\eta) \cap [z - \eta, z + \eta]$ and note that $\rho \in P(x,z)$ since $\Theta(\rho) = (-\infty, \theta') \cup (\theta'', +\infty)$, where θ' and θ'' solve $\theta' = C(\rho) = \Psi((x^* - \theta)/\sigma)$ and satisfy $0 < \theta' < \theta^* < \tilde{\theta} < \theta^{**} < \theta'' < \infty$, which implies that $[\theta^*, \theta^{**}] \cap \Theta(\rho) = \varnothing$. Then, take any μ that satisfies $\mu(\{r = \rho\}|x,z) = 1$, $\mu(\{\theta \in \Theta(x) \cap \Theta(\rho), r = \rho\}|x,z) = 0$ and $\mu(\{\theta <$

$\hat{\theta}, r=\rho\}|x,z) > \rho$ if and only if $x < x^*$. If instead $\Theta(x) \cap [\theta^*, \theta^{**}] = \varnothing$,then either $P(x,z) = \varnothing$, in which case simply take any $\rho \in [z-\eta, r^*-2\eta]$ and any μ such that $\mu(\{r=\rho\}|x,z) = 1$, or $P(x,z) \neq \varnothing$, in which case there must exist a $\rho \in (z-\eta, r^*-2\eta)$ such that $\rho \in P(x,z)$. Then take any μ such that $\mu(\{r=\rho\}|x,z) = 1$ and $\mu(\{\theta \in \Theta(x) \cap \Theta(\rho), r=\rho\}|x,z) = 0$. Finally note that $\Theta(x) \cap [\theta^*, \theta^{**}] \neq \varnothing$ implies that ξ is necessarily bounded and hence either $x < \theta^* - \sigma < x^*$ in which case $\mu(\{\theta < \hat{\theta}, r=\rho\}|x,z) = 1$, or $x > \theta^{**} + \sigma > x^*$ in which case $\mu(\{\theta < \hat{\theta}, r=\rho\}|x,z) = 0$.

When instead $z \in [r^*-\eta, r^*+\eta]$ and $\Theta(x) \cap [\theta^*, \theta^{**}] \neq \varnothing$, beliefs are again pinned down by Bayes' rule and satisfy $\mu(\{\theta \in [\theta^*, \theta^{**}], r=r^*\}|x,z) = 1$. When instead $z \in [r^*-\eta, r^*+\eta]$ and $\Theta(x) \cap [\theta^*, \theta^{**}] = \varnothing$, then take any $\rho \in [r^*, z+\eta] \cap \mathcal{R}$ and any μ such that $\mu(\{r=\rho\}|x,z) = 1$ and $\mu(\{\theta < 0, r=\rho\}|x,z) = 0$ for any $x \geq \underline{x}$ so that attacking if and only if $x < \underline{x}$ is optimal. Note that $\rho > r^*$ is dominated in equilibrium for all $\theta \in \Theta(x)$. Nevertheless, in this case, we do not need to restrict beliefs to assign positive measure only to $r \in P(x,z)$ and $\theta \notin \Theta(x) \cap \Theta(r)$, for such restrictions would make the agents (weakly) more aggressive, thus making the deviation even less profitable for the policy maker.

Finally, for any $z > r^*+\eta$, since necessarily $P(x,z) = \varnothing$, simply take any $\rho \in [z-\eta, z+\eta] \cap \mathcal{R}$ and any μ such that $\mu(\{r=\rho\}|x,z) = 1$ and $\mu(\{\theta < 0, r=\rho\}|x,z) > 0$ if and only if $x < \underline{x}$.

Given these beliefs, the strategy of the agents is sequentially rational for any (x,z).

Next, consider the policy maker. For $\theta < 0$, $\underline{r}$ is dominant. For $\theta \geq 0$,

$$A(\theta, r) = \begin{cases} 0 & \text{for } r \geq r^* \\ \Psi(\frac{x^*-\theta}{\sigma})\Phi(\frac{z^*-r}{\eta}) & \text{for } r \in [z^*-\eta, r^*) \\ \Psi(\frac{x^*-\theta}{\sigma}) & \text{for } r \leq z^*-\eta \end{cases}$$

Clearly, $\underline{r}$ is preferred to any $r \in (\underline{r}, z^*-\eta)$ and r^* is preferred to any $r > r^*$ by all θ. The payoff associated to r^* is $\theta - C(r^*)$, while the payoff associated to any $r \in [z^*-\eta, r^*)$ is $\max\{\theta - A(\theta, r), 0\} - C(r)$. Hence, r^* is preferred to $r \in [z^*-\eta, r^*)$ if and only if $C(r^*) - C(r) \leq \min\{A(\theta, r), \theta\}$.

For any $\theta \in [\theta^*, \theta^{**}]$, since $\theta^* = C(r^*)$, $C(r^*) - C(r) < \theta$, implying that r^* is preferred to $r \in [z^*-\eta, r^*)$ if and only if $C(r^*) - C(r) \leq A(\theta, r)$. Furthermore, since $C(r^*) = A(\theta^{**}, \underline{r}) = \Psi(\frac{x^*-\theta^{**}}{\sigma})$, this is satisfied for all $\theta \in [\theta^*, \theta^{**}]$ if and only if

$$C(r^*) - C(r) \leq C(r^*)\,\Phi(\tfrac{z^*-r}{\eta}). \tag{5}$$

By Lipschitz continuity of C and absolute continuity of Φ, $C(r^*) - C(r) \leq K(r^*-r)$ and $\Phi(\frac{z^*-r}{\eta}) = \int_{-1}^{(z^*-r)/\eta} \phi(\zeta)d\zeta \geq \frac{1}{\eta}\underline{\phi}[r^*-r]$, whereas $r^* > r_\eta$ implies $K < C(r^*)\underline{\phi}/\eta$. It follows that, for all

$r \in [z^* - \eta, r^*)$,

$$C(r^*) - C(r) - C(r^*)\,\Phi(\tfrac{z^*-r}{\eta}) \leq [K - \tfrac{1}{\eta}\underline{\phi} C(r^*)][r^* - r] < 0, \tag{6}$$

which in turn suffices for (5). Furthermore, since $\theta - C(r^*) \geq \max\{\theta - \Psi(\frac{x^*-\theta}{\sigma}), 0\}$ for all $\theta \in [\theta^*, \theta^{**}]$, r^* is also optimal for all $\theta \in [\theta^*, \theta^{**}]$.

Next, consider $\theta \in [0, \theta^*)$. In this case, $\Psi(\frac{x^*-\theta}{\sigma}) > \theta^* > \theta$ and therefore $\min\{A(\theta, r), \theta\} \geq \theta\Phi(\frac{z^*-r}{\eta})$, so that the payoff from setting $r \in [z^* - \eta, r^*]$ is

$$\theta - \min\{A(\theta, r), \theta\} - C(r) \leq \theta[1 - \Phi(\tfrac{z^*-r}{\eta})] - C(r).$$

Hence, for $\underline{r}$ to be optimal for any $\theta \in [0, \theta^*)$, it suffices that $\theta[1 - \Phi(\frac{z^*-r}{\eta})] - C(r) \leq 0$. But this follows immediately from (6) using $\theta < \theta^* = C(r^*)$.

Finally, consider $\theta > \theta^{**}$. In this case, $\Psi(\frac{x^*-\theta}{\sigma}) \leq C(r^*) < \theta$ and therefore the payoff from setting $r \in [z^* - \eta, r^*]$ is smaller than the payoff associated with $\underline{r}$ if $\theta - A(\theta, r) - C(r) \leq \theta - \Psi(\frac{x^*-\theta}{\sigma})$, or equivalently $\Psi(\frac{x^*-\theta}{\sigma})[1 - \Phi(\frac{z^*-r}{\eta})] - C(r) \leq 0$. Using $\Psi(\frac{x^*-\theta^{**}}{\sigma}) = C(r^*)$, this is satisfied for all $\theta > \theta^{**}$ if and only if $C(r^*) - C(r^*)\,\Phi(\frac{z^*-r}{\eta}) - C(r) \leq 0$ for all $r \in [z^* - \eta, r^*)$, which once again follows by (6) when $r^* > r_\eta$. *QED*

Proof of Proposition 5. *Part (i).* When $r(\theta) = \underline{r}$ for all θ, z conveys no information about θ and hence $\Pr[D = 1|x, z] = \Pr[\theta \leq \tilde{\theta}|x, z] = 1 - \Psi(\frac{x-\tilde{\theta}}{\sigma})$ and $\mathbb{E}[r|x, z] = \underline{r}$ for any z. An agent thus finds it optimal to attack if and only if $x < \tilde{x}$, where $\tilde{x}$ solves $\Psi(\frac{\tilde{x}-\tilde{\theta}}{\sigma}) = 1 - \underline{r}$. The size of the attack is then given by $A(\theta, r) = \Psi(\frac{\tilde{x}-\theta}{\sigma})$ and is independent of r, implying that the policy maker indeed finds it optimal to set $r(\theta) = \underline{r}$ for all θ and abandon the regime if and only if $\theta < \tilde{\theta}$, where $\tilde{\theta}$ solves $\Psi(\frac{\tilde{x}-\tilde{\theta}}{\sigma}) = \tilde{\theta}$.

Part (ii). Let $r^* \in (\underline{r}, \tilde{r})$. We prove that, for η small enough, there exist thresholds $(\theta', \theta'', x', \hat{x})$ and an equilibrium such that the policy maker sets $r(\theta) = r^*$ if $\theta \in [\theta', \theta'']$ and $r(\theta) = \underline{r}$ otherwise, an agent attacks if and only if either $x < \hat{x}$ or $(x, z) < (x', z^*)$, where $z^* = r^*$, and the status quo is abandoned if and only if $\theta < \theta'$.

Consider first the agents. Since ζ is exponential (i.e., $\Phi(\zeta) = 1 - \exp(-\zeta)$), the likelihood ratio of r^* vs $\underline{r}$ conditional on z is $\rho + (1 - \rho)\exp(\frac{r^*-\underline{r}}{\eta})$ for $z \geq z^*$ and ρ for $z < z^*$. The expected payoff from attacking for an agent with signals (x,z), $u(x,z) \equiv \Pr\left[\theta \leq \theta' | x,z\right] - \mathbb{E}\left[r|x,z\right]$, is thus

$$u(x,z) = \begin{cases} \dfrac{1-\Psi(\frac{x-\theta'}{\sigma}) - \{\underline{r} - (\underline{r} - \rho r^*)[\Psi(\frac{x-\theta'}{\sigma}) - \Psi(\frac{x-\theta''}{\sigma})]\}}{1-(1-\rho)[\Psi(\frac{x-\theta'}{\sigma}) - \Psi(\frac{x-\theta''}{\sigma})]} & \text{for } z < r^* \\ \dfrac{1-\Psi(\frac{x-\theta'}{\sigma}) - \{\underline{r} + [r^*\rho + r^*(1-\rho)\exp(\frac{r^*-r}{\eta}) - \underline{r}][\Psi(\frac{x-\theta'}{\sigma}) - \Psi(\frac{x-\theta''}{\sigma})]\}}{1+[\rho + (1-\rho)\exp(\frac{r^*-r}{\eta}) - 1][\Psi(\frac{x-\theta'}{\sigma}) - \Psi(\frac{x-\theta''}{\sigma})]} & \text{for } z \geq r^* \end{cases} \tag{7}$$

Note that, for any z, u is continuous in x, $u \to 1-\underline{r} > 0$ as $x \to -\infty$ and $u \to -\underline{r} < 0$ as $x \to +\infty$. Furthermore, for any x, $u(x,z)$ is a step function in z, with discontinuity at $z = r^*$. It follows that there exist thresholds x' and $\hat{x}$ such that x' solves $u(x',z) = 0$ for $z < r^*$ and $\hat{x}$ solves $u(\hat{x},z) = 0$ for $z \geq r^*$; equivalently,

$$1-\Psi(\tfrac{x'-\theta'}{\sigma}) = \underline{r} - (\underline{r} - \rho r^*)[\Psi(\tfrac{x'-\theta'}{\sigma}) - \Psi(\tfrac{x'-\theta''}{\sigma})] \tag{8}$$

$$1-\Psi(\tfrac{\hat{x}-\theta'}{\sigma}) = \underline{r} + [r^*\rho + r^*(1-\rho)\exp(\tfrac{r^*-r}{\eta}) - \underline{r}][\Psi(\tfrac{\hat{x}-\theta'}{\sigma}) - \Psi(\tfrac{\hat{x}-\theta''}{\sigma})]. \tag{9}$$

Next, note that $u(x,z) = N(x,z)/D(x,z)$, where $N(x,z)$ and $D(x,z) > 0$ are respectively the numerator and the denominator in (7). When $z < r^*$, assuming $\rho < \underline{r}/r^*$ suffices for $N(x,z)$ to be strictly decreasing in x and therefore for $\partial u(x',z)/\partial x < 0$. For $z \geq r^*$, on the other hand, note that $N(x,z) = [1-\Psi(\frac{x-\theta'}{\sigma})]H(x)$, where

$$H(x) = 1 - \frac{\underline{r}}{1-\Psi(\frac{x-\theta'}{\sigma})} - \frac{[r^*\rho + r^*(1-\rho)\exp(\frac{r^*-r}{\eta}) - \underline{r}][\Psi(\frac{x-\theta'}{\sigma}) - \Psi(\frac{x-\theta''}{\sigma})]}{1-\Psi(\frac{x-\theta'}{\sigma})}.$$

At $x = \hat{x}$, necessarily $H(\hat{x}) = 0$, which implies that $\partial N(\hat{x},z)/\partial x = H'(\hat{x}) < 0$ since ψ and (hence $1-\Psi$) is log-concave, and therefore $\partial u(\hat{x},z)/\partial x < 0$. It follows that, given θ' and θ'', x' and $\hat{x}$ are the *unique* solutions to (8)-(9).

Consider now the behavior of the policy maker. When agents follow the strategies described above, the size of attack is given by

$$A(\theta,r) = \begin{cases} [1-\rho\exp(-\frac{r^*-r}{\eta}) - (1-\rho)\exp(-\frac{r^*-\underline{r}}{\eta})]\Psi(\frac{x'-\theta}{\sigma}) + \\ +[\rho\exp(-\frac{r^*-r}{\eta}) + (1-\rho)\exp(-\frac{r^*-\underline{r}}{\eta})]\Psi(\frac{\hat{x}-\theta}{\sigma}) & \text{for } r < r^* \\ \rho[1-\exp(-\frac{r^*-\underline{r}}{\eta})]\Psi(\frac{x'-\theta}{\sigma}) + [1-\rho+\rho\exp(-\frac{r^*-\underline{r}}{\eta})]\Psi(\frac{\hat{x}-\theta}{\sigma}) & \text{for } r \geq r^* \end{cases}$$

$A(\theta,r)$ is strictly decreasing in θ, equals $A(\theta,r^*)$ for any $r \geq r^*$, and is strictly decreasing and strictly concave in $r \in (\underline{r}, r^*)$. Together with the linearity of C, this implies that any $r \notin \{\underline{r}, r^*\}$ is dominated by either $\underline{r}$ or r^*. For the proposed strategy to be optimal, it must be that the policy maker prefers to set $r = \underline{r}$ and abandon the status quo for $\theta < \theta'$, set $r = r^*$ and maintain for $\theta \in [\theta', \theta'']$, and maintain while setting $r = \underline{r}$ for $\theta > \theta''$.

Let $U_1(\theta)$ and $U_2(\theta)$ denote the payoffs from setting, respectively, r^* and $\underline{r}$ while maintaining the status quo:

$$\begin{aligned} U_1(\theta) &\equiv \theta - \rho[1 - \exp(-\tfrac{r^* - \underline{r}}{\eta})]\Psi(\tfrac{x' - \theta}{\sigma}) - [1 - \rho + \rho \exp(-\tfrac{r^* - \underline{r}}{\eta})]\Psi(\tfrac{\hat{x} - \theta}{\sigma}) - C(r^*) \\ U_2(\theta) &\equiv \theta - [1 - \exp(-\tfrac{r^* - \underline{r}}{\eta})]\Psi(\tfrac{x' - \theta}{\sigma}) - \exp(-\tfrac{r^* - \underline{r}}{\eta})\Psi(\tfrac{\hat{x} - \theta}{\sigma}). \end{aligned}$$

The two thresholds θ' and θ'' must thus solve $U_1(\theta') = 0$ and $U_2(\theta'') = U_1(\theta'')$, or equivalently

$$\theta' = \rho[1 - \exp(-\tfrac{r^* - \underline{r}}{\eta})]\Psi(\tfrac{x' - \theta'}{\sigma}) + [1 - \rho + \rho \exp(-\tfrac{r^* - \underline{r}}{\eta})]\Psi(\tfrac{\hat{x} - \theta'}{\sigma}) + C(r^*) \quad (10)$$

$$C(r^*) = (1 - \rho)[1 - \exp(-\tfrac{r^* - \underline{r}}{\eta})][\Psi(\tfrac{x' - \theta''}{\sigma}) - \Psi(\tfrac{\hat{x} - \theta''}{\sigma})]. \quad (11)$$

Let $q(\theta) \equiv (1 - \rho)[1 - \exp(-\frac{r^* - \underline{r}}{\eta})][\Psi(\frac{x' - \theta}{\sigma}) - \Psi(\frac{\hat{x} - \theta}{\sigma})]$. Since $\hat{x} < x'$, the distribution $1 - \Psi(\frac{x' - \theta}{\sigma})$ first order stochastically dominates the distribution $1 - \Psi(\frac{\hat{x} - \theta}{\sigma})$. It follows that $\psi(\frac{x' - \theta}{\sigma})/\psi(\frac{\hat{x} - \theta}{\sigma})$ is increasing in θ, implying that there exists a unique $\hat{\theta}$ such that $\psi(\frac{x' - \theta}{\sigma}) \lesseqqgtr \psi(\frac{\hat{x} - \theta}{\sigma})$ if and only if $\theta \lesseqqgtr \hat{\theta}$; equivalently, $q(\theta)$ is increasing for $\theta < \hat{\theta}$ and decreasing for $\theta > \hat{\theta}$. Furthermore, $\lim_{\theta \to -\infty} q(\theta) = \lim_{\theta \to +\infty} q(\theta) = 0$ and therefore (11) admits at most two solutions.

To sustain the proposed equilibrium, θ' must be between the two solutions, θ_1 and θ_2, of (11). Indeed, provided that $\theta' \in [\theta_1, \theta_2]$, $U_1(\theta) \geq U_2(\theta)$ if and only if $\theta \in [\theta_1, \theta_2]$ and $U_1(\theta) \geq 0$ if and only if $\theta \geq \theta'$; since U_1 and U_2 are increasing in θ, the strategy for the policy maker is then optimal with $\theta'' = \theta_2$. If instead $\theta' < \theta_1$, setting r^* would not be optimal for $\theta \in [\theta', \theta_1)$; and if $\theta' > \theta_2$, r^* would never be optimal. Finally, note that $\theta' \in [\theta_1, \theta_2]$ if and only if

$$C(r^*) \leq (1 - \rho)[1 - \exp(-\tfrac{r^* - \underline{r}}{\eta})][\Psi(\tfrac{x' - \theta'}{\sigma}) - \Psi(\tfrac{\hat{x} - \theta'}{\sigma})]. \quad (12)$$

The following lemma completes the proof by showing that, for η small enough, the proposed equilibrium exists and is close to the corresponding one in the game without policy noise.

Lemma. *For any $r^* \in (\underline{r}, \tilde{r})$ and $\varepsilon > 0$, there exist $\bar{\eta} > 0$ and $\bar{\rho} < \underline{r}/r^*$ such that for any $(\eta, \rho) < (\bar{\eta}, \bar{\rho})$, equations (8), (9), (10) and (11) admit a solution $(x', \hat{x}, \theta', \theta'')$ that satisfies (12), $\theta' \leq \theta''$, $|x' - x^*| < \varepsilon$, $|\theta' - \theta^*| < \varepsilon$, $|\theta'' - \theta^{**}| < \varepsilon$ and $\hat{x} < -1/\varepsilon$.*

Proof. Let

$$W = \Psi(\tfrac{\hat{x}-\theta'}{\sigma}), \qquad Z = \Psi(\tfrac{\hat{x}-\theta''}{\sigma}), \qquad Y = \Psi(\tfrac{x'-\theta''}{\sigma}), \tag{13}$$

Conditions (8)-(11) can then be rewritten as follows:

$$\delta - \gamma Y = \Psi\left(\Psi^{-1}(Y) - \Psi^{-1}(Z) + \Psi^{-1}(W)\right) \tag{14}$$

$$W = \alpha + \beta Z \tag{15}$$

$$\theta' = \rho[1 - \exp(-\tfrac{r^* - \underline{r}}{\eta})]\left[\Psi\left(\Psi^{-1}(Y) - \Psi^{-1}(Z) + \Psi^{-1}(W)\right) - W\right] + W + C(r^*) \tag{16}$$

$$Y = Z + \frac{C(r^*)}{(1-\rho)[1 - \exp(-\frac{r^*-\underline{r}}{\eta})]} \tag{17}$$

where

$$\alpha \equiv \tfrac{1-\underline{r}}{1-\underline{r}+r^*[\rho+(1-\rho)\exp(\frac{r^*-\underline{r}}{\eta})]}, \quad \beta \equiv \tfrac{r^*[\rho+(1-\rho)\exp(\frac{r^*-\underline{r}}{\eta})]-\underline{r}}{1-\underline{r}+r^*[\rho+(1-\rho)\exp(\frac{r^*-\underline{r}}{\eta})]}, \quad \gamma \equiv \tfrac{\underline{r}-\rho r^*}{1-\underline{r}+\rho r^*}, \quad \delta \equiv \tfrac{1-\underline{r}}{1-\underline{r}+\rho r^*}.$$

Note that $\alpha, \beta, \delta \in (0,1)$ and $\gamma > 0$. Substituting (15) into (14) gives

$$\Psi^{-1}(\delta - \gamma Y) - \Psi^{-1}(Y) = \Psi^{-1}(\alpha + \beta Z) - \Psi^{-1}(Z). \tag{18}$$

Let $LHS(Y)$ and $RHS(Z)$ denote, respectively, the left-hand and the right-hand side of (18). Note that $LHS(Y)$ and $RHS(Z)$ are defined for $Y \in (0, \min\{1, \delta/\gamma\})$ and $Z \in (0,1)$ and are continuous in Y and Z. Moreover, LHS is decreasing in Y, with $\lim_{Y\to 0} LHS(Y) = \infty$, $\lim_{Y\to\min\{1,\delta/\gamma\}} LHS(Y) = -\infty$ and $LHS(Y) \gtreqless 0$ if and only if $Y \lesseqgtr 1 - \underline{r}$, whereas $\lim_{Z\to 0} RHS(Z) = \infty$, $\lim_{Z\to 1} RHS(Z) = -\infty$, and $RHS(Z) \gtreqless 0$ if and only if $Z \lesseqgtr 1 - \underline{r}$. It follows that (18) defines implicitly a continuous function $Y = g(Z;\eta,\rho)$, with $g : (0,1) \times \mathbb{R}^2 \to (0, \min\{1, \delta/\gamma\})$; note that $\lim_{Z\to 0} g(Z) = 0$, $\lim_{Z\to 1} g(Z) = \min\{1, \delta/\gamma\}$, and $g(Z) \lesseqgtr 1 - \underline{r}$ if and only if $Z \lesseqgtr 1 - \underline{r}$. Condition (17), on the other hand, defines explicitly a function $Y = f(Z;\eta,\rho)$.

We want to prove that (17) and (18), or equivalently $Y = f(Z) = g(Z)$, admit a solution for (Y,Z). Note that $f(Z;\eta,\rho)$ is continuous and increasing in (Z,η,ρ) with $f(0;\eta,\rho) \to C(r^*) \in (0, 1-\underline{r})$ as $(\eta, r) \to (0,0)$. Then, take any (η_0, ρ_0, Z_0) such that $f(Z_0;\eta_0,\rho_0) < 1-\underline{r}$, and note that $g(Z;\eta,\rho)$ is also continuous in (Z,η,ρ) with $g(Z_0;\eta,\rho) \to 1-\underline{r}$ as $\eta \to 0$ and $g(Z;\eta,\rho) \to 0$ for any (η,ρ) as $Z \to 0$. It follows that there exist $\tilde{\eta} \in (0,\eta_0)$, $\tilde{\rho} < \min\{\rho_0, \underline{r}/r^*\}$ and $Z_1 < Z_0$ such that for any $(\eta,\rho) < (\tilde{\eta},\tilde{\rho})$, $g(Z_0;\eta,\rho) > f(Z_0;\eta,\rho)$ and $g(Z_1;\eta,\rho) < f(Z_1;\eta,\rho)$. The graphs of g and f thus intersect at least twice for (η,ρ) sufficiently small, implying that the system $Y = f(Z) = g(Z)$ admits at least two solutions, as illustrated in Figure A1.

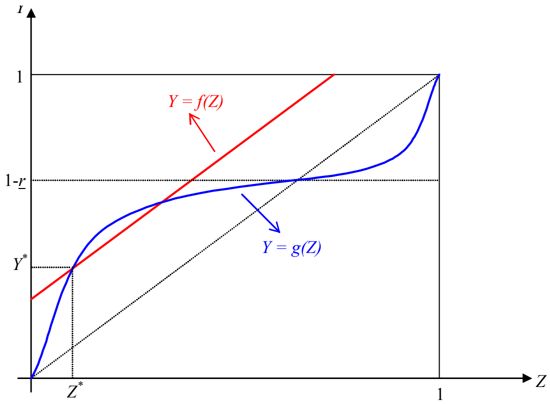

Figure A1

Consider the lowest solution (Z^*, Y^*), let $W^* = \alpha + \beta Z^*$ and note that (Z^*, Y^*, W^*) are continuous in (η, ρ) and satisfy $Z^* \in (0, 1 - \underline{r})$, $Y^* \in (Z^*, 1 - \underline{r})$ and $W^* \in (Z^*, 1 - \underline{r})$. The thresholds $(x', \hat{x}, \theta', \theta'')$ are then the unique solutions to (13) and (16). That $W^* > Z^*$ and $Y^* > Z^*$ imply that $0 < \theta' < \theta''$. It remains to show that for η and ρ sufficiently small (12) is also satisfied, in which case $U_2(\theta') < U_1(\theta') = 0$ and hence $\theta' < A(\theta', \underline{r}) < 1$. Using $\Psi(\frac{x'-\theta'}{\sigma}) = \Psi(\Psi^{-1}(Y^*) - \Psi^{-1}(Z^*) + \Psi^{-1}(W^*))$ and (14), we then have that $\Psi(\frac{x'-\theta'}{\sigma}) = \delta - \gamma Y^*$ and hence $\Psi(\frac{x'-\theta'}{\sigma}) - \Psi(\frac{\hat{x}-\theta'}{\sigma}) = \delta - \gamma Y^* - W^*$. Moreover, as $(\eta, \rho) \to (0, 0)$, $Z^* \to 0$, $Y^* \to C(r^*)$, $W^* \to 0$, $\delta \to 1$, $\gamma \to \frac{\underline{r}}{1-\underline{r}}$ and $\exp(-\frac{r^*-\underline{r}}{\eta}) \to 0$, implying that

$$(1-\rho)[1-\exp(-\tfrac{r^*-\underline{r}}{\eta})][\Psi(\tfrac{x'-\theta'}{\sigma}) - \Psi(\tfrac{\hat{x}-\theta'}{\sigma})] \to 1 - \tfrac{\underline{r}}{1-\underline{r}} C(r^*).$$

But since $C(r^*) < 1 - \underline{r}$, then necessarily $1 - \frac{\underline{r}}{1-\underline{r}} C(r^*) > 1 - \underline{r} > C(r^*)$, which implies that there exists $\eta' \in (0, \tilde{\eta})$ and $\rho' \in (0, \tilde{\rho})$, such that for any $(\eta, \rho) < (\eta', \rho')$, the solution to (8), (9), (10) and (11) satisfies (12).

Finally, note that as $(\eta, \rho) \to (0, 0)$, $Y^* \to C(r^*)$, $W^* \to 0$, and $Z^* \to 0$. Using (13), (16) and (8), we then have that $\theta' \to C(r^*)$, $\hat{x} = \theta' + \sigma \Psi^{-1}(W) \to -\infty$, $x' \to x^*$, and $\theta'' \to \theta^{**}$. Hence, for any $\varepsilon > 0$, there exist $\bar{\eta} \in (0, \eta')$ and $\bar{\rho} \in (0, \rho')$ such that $(\eta, \rho) < (\bar{\eta}, \bar{\rho})$ suffices for $|x' - x^*| < \varepsilon$, $|\theta' - \theta^*| < \varepsilon$, $|\theta'' - \theta^{**}| < \varepsilon$ and $\hat{x} < -1/\varepsilon$, where $(x^*, \theta^*, \theta^{**})$ are defined as in Proposition (3). □ *QED*

Proof of Proposition 6. We prove the result in three steps. The construction of an equilibrium in which the policy is raised if and only if $\theta \in \Theta^{**}$ is in step 3; steps 1 and 2 characterize the set Θ^{**} and establish conditions that are useful for step 3.

Step 1. Fix $r^* \in (\underline{r}, \tilde{r})$ and let $\theta^* = \min\{\theta : R(\theta, 0, r^*) \geq 0 \text{ and } U(\theta, 0, r^*) \geq 0\} = \min\{\theta \geq 0 : U(\theta, 0, r^*) \geq 0\}$. Next, define the function $m : \mathbb{R}^2 \to [0,1]$ and the correspondence $S : \mathbb{R} \to 2^{\mathbb{R}_+}$ as follows:

$$\begin{aligned} m\left(x, x'\right) &\equiv \frac{1 - \Psi(\frac{x'-\theta^*}{\sigma})}{1 - \int_{S(x)} \frac{1}{\sigma}\psi(\frac{x'-\theta^*}{\sigma})d\theta}, \\ S(x) &\equiv \left\{\theta \geq \theta^* : R\left(\theta, \Psi\left(\tfrac{x-\theta}{\sigma}\right), \underline{r}\right) < 0 \text{ or } U\left(\theta, \Psi\left(\tfrac{x-\theta}{\sigma}\right), \underline{r}\right) \leq U(\theta, 0, r^*)\right\}. \end{aligned}$$

Step 2 below shows that either there exists an $x^* \in \mathbb{R}$ such that $m(x^*, x^*) = g(\underline{r})$, or $m(x, x) > g(\underline{r})$ for all x, in which case we let $x^* = \infty$. In either case, let $\Theta^{**} = S(x^*)$.

θ^* is the lowest type who is willing to raise the policy at r^* if this ensures that no agent attacks. $S(x)$ is the set of $\theta \geq \theta^*$ who prefer r^* to $\underline{r}$ when agents do not attack when $r = r^*$ and attack if and only if their signal is less than x when $r = \underline{r}$. $m(x, x')$ in turn is the posterior probability of regime change for an agent with signal x' when he observes $\underline{r}$ and believes that the regime is abandoned if and only if $\theta < \theta^*$ and that the policy is $r(\theta) = \underline{r}$ if and only if $\theta \notin S(x)$. The triplet $(x^*, \theta^*, \Theta^{**})$ thus identify an equilibrium for the fictitious game in which the policy maker is restricted to set $r \in \{\underline{r}, r^*\}$ and the agents are restricted not to attack when $r = r^*$. Step 3 shows that this is also part of an equilibrium for the unrestricted game.

Step 2. Note that $S(x)$ is continuous in x and $S(x_1) \subseteq S(x_2)$ for any $x_1 \leq x_2$ (because $R(\theta, A, r)$ and $U(\theta, A, r)$ are non-increasing and continuous in A and $\Psi\left(\frac{x-\theta}{\sigma}\right)$ is non-increasing and continuous in x), whereas $m(x, x')$ is continuous in (x, x'), non-decreasing in x (by the monotonicity of S) and non-increasing in x' (by the log-concavity of ψ). Moreover, for any $(x, x') \in \mathbb{R}^2$, we have $S(x) \subseteq \bar{S} \equiv \{\theta \geq \theta^* : R(\theta, 1, \underline{r}) < 0 \text{ or } U(\theta, 1, \underline{r}) \leq U(\theta, 0, r^*)\}$ and

$$1 \geq \frac{1 - \Psi(\frac{x'-\theta^*}{\sigma})}{1 - \int_{\bar{S}} \frac{1}{\sigma}\psi(\frac{x'-\theta^*}{\sigma})d\theta} \geq m\left(x, x'\right) \geq 1 - \Psi(\tfrac{x'-\theta^*}{\sigma}).$$

It follows that, for any x, $m(x, x') \geq g(\underline{r})$ for all $x' \leq x^{\#}$, where $x^{\#} \in \mathbb{R}$ is the solution to $1 - \Psi\left(\left(x^{\#} - \theta^*\right)/\sigma\right) = g(\underline{r})$.

Define now the sequence $\{x_k\}_{k=0}^{\infty}$, with $x_k \in \mathbb{R} \cup \{+\infty\}$, as follows: for $k = 0$, let $x_0 = x^{\#}$; for $k \geq 1$, let x_k be the solution to $m(x_{k-1}, x_k) = g(\underline{r})$ if $x_{k-1} < \infty$ and $\inf\{x' : m(x_{k-1}, x') \leq g(\underline{r})\} < \infty$, and $x_k = \infty$ otherwise. The fact that $m\left(x^{\#}, x^{\#}\right) \geq 1 - \Psi(\frac{x^{\#}-\theta^*}{\sigma}) = g(\underline{r})$, together with the continuity and monotonicities of m, ensures that this sequence is well defined and non-decreasing. It follows that either $\lim_{k\to\infty} x_k \in [x^{\#}, +\infty)$, or $\lim_{k\to\infty} x_k = +\infty$. In the former case, let $x^* = \lim x_k$ and $\Theta^{**} = S(x^*)$; in the latter, let $x^* = \infty$ and $\Theta^{**} = S(\infty) \equiv \bar{S}$.

Note that $\theta^* \in [0, \tilde{\theta})$ and $x^* > \hat{x}$, where $\hat{x} \in \mathbb{R}$ is the solution to $R(\theta^*, \Psi((\hat{x} - \theta^*)/\sigma), \underline{r}) = 0$. That $\theta^* < \tilde{\theta}$ follows immediately from $r^* < \tilde{r}$ and hence $U(\tilde{\theta}, 0, r^*) > U(\tilde{\theta}, 0, \tilde{r}) = 0$. To see that $x^* > \hat{x}$, note that, by the definitions of $\hat{x}$, $\tilde{\theta}$, and $x^{\#}$, $R(\theta^*, \Psi(\frac{x^{\#}-\theta^*}{\sigma}), \underline{r}) = 0 = R(\tilde{\theta}, 1 - g(\underline{r}), \underline{r}) = R(\tilde{\theta}, \Psi(\frac{x^{\#}-\theta^*}{\sigma}), \underline{r})$, which together with $\theta^* < \tilde{\theta}$ implies $x^{\#} > \hat{x}$ and therefore $x^* \geq x^{\#} > \hat{x}$. This in turn implies that there exists a $\hat{\theta} \in (\theta^*, 1)$ which solves $R(\hat{\theta}, \Psi(\frac{x^{\#}-\theta^*}{\sigma}), \underline{r}) = 0$ such that $R(\theta, \Psi(\frac{x^{\#}-\theta^*}{\sigma}), \underline{r}) < 0$ if and only if $\theta < \hat{\theta}$. But then $[\theta^*, \hat{\theta}) \subseteq S(x^*)$.

Finally, note that, when the noise is bounded, $m(x, x') = 0$ for all $x' \geq \theta^* + \sigma$, and $S(x) \subseteq [\theta^*, x + \sigma]$ if $x + \sigma \geq \theta^*$ and $S(x) = \varnothing$ otherwise. It follows that, with bounded noise, $x^* < \theta^* + \sigma$, $\Theta^{**} \subseteq [\theta^*, \theta^* + 2\sigma]$, and $\hat{\theta} \in (\theta^*, \theta^* + 2\sigma)$.

Step 3. Define the function $X : [\underline{r}, \overline{r}] \to \mathbb{R} \cup \{\pm\infty\}$ as follows: at $r = \underline{r}$, let $X(\underline{r}) = x^*$; for $r \in (\underline{r}, r^*)$, let $X(r) = \infty$ if $x^* = \infty$ and otherwise let $X(r) \geq x^*$ be the solution to $R(\hat{\theta}, \Psi(\frac{X(r)-\hat{\theta}}{\sigma}), r) = 0$; finally, for $r \in [r^*, \overline{r}]$, let $X(r) = \underline{x}$. We now show that the following strategies are part of an equilibrium: the policy maker sets $r(\theta) = r^*$ for $\theta \in \Theta^{**}$ and $r(\theta) = \underline{r}$ otherwise; an agent attacks if and only if $x < X(r)$.

Consider first the policy maker. By construction of $X(r)$, for any $r < r^*$, $A(\theta, r) \geq A(\theta, \underline{r})$ and $sign\{R(\theta, A(\theta, r), r)\} = sign\{R(\theta, A(\theta, \underline{r}), \underline{r})\}$, whereas for any $r \geq r^*$, $A(\theta, r) = A(\theta, r^*)$ and $sign\{R(\theta, A(\theta, r), r)\} = sign\{R(\theta, A(\theta, r^*), r^*)\}$. It follows that the policy maker strictly prefers $\underline{r}$ to any $r \in (\underline{r}, r^*)$ and r^* to any $r > r^*$. For any $\theta < \theta^*$, $R(\theta, A(\theta, \underline{r}), \underline{r}) < 0$ and $U(\theta, 0, r^*) < 0$, which implies that any $\theta < \theta^*$ finds it optimal to set $\underline{r}$ and then face regime change. On the contrary, any $\theta > \theta^*$, necessarily maintains the status quo, since setting r^* guarantees that $R(\theta, 0, r^*) > 0$ and $U(\theta, 0, r^*) > 0$. By definition of S, $\Theta^{**} = S(x^*)$ is then the set of types above θ^* who prefer to raise the policy at r^* than setting $\underline{r}$. We thus conclude that $r(\theta) = r^*$ if $\theta \in \Theta^{**}$ and $r(\theta) = \underline{r}$ otherwise is indeed optimal for the policy maker.

Consider next the agents. Given $r = \underline{r}$, beliefs are necessarily pinned down by Bayes rule since any x is consistent with either $(-\infty, \theta^*)$ or $(\theta^*, +\infty) \backslash \Theta^{**}$; this is immediate in the case of unbounded noise and is ensured by the fact that $\Theta^{**} \subseteq [\theta^*, \theta^* + 2\sigma]$ in the case of bounded noise. The posterior probability of regime change is then given by

$$\mu(\theta^* | x, \underline{r}) = \frac{1 - \Psi(\frac{x-\theta^*}{\sigma})}{1 - \int_{\Theta^{**}} \frac{1}{\sigma} \psi(\frac{x-\theta^*}{\sigma}) d\theta} = m(x^*, x),$$

and is decreasing in x. Moreover, by definition of x^*, either $x^* < +\infty$ and $m(x^*, x^*) = g(\underline{r})$, or $x^* = +\infty$ in which case the probability of regime change is $m(x^*, x) \geq g(\underline{r})$ for all x. Hence, given $\underline{r}$, it is indeed optimal to attack if and only if $x < x^*$. When instead $r = r^*$, Bayes's rule implies

$\mu(\theta^*|x,r^*) = 0$ for any x such that $\Theta(x) \cap \Theta^{**} \neq \varnothing$, in which case it is optimal not to attack.

For out-of-equilibrium events, we follow a construction similar to that in Proposition 3. The set of types for whom a deviation $r \notin \{\underline{r}, r^*\}$ is dominated in equilibrium is $\Theta(r) = (-\infty, 0] \cup \{\theta \geq 0 : U(\theta, 0, r) < U^*(\theta)\}$, where $U^*(\theta) = \max\{0, U(\theta, A(\theta, r(\theta)), r(\theta))\}$ denotes the equilibrium payoff. For any $r \in (\underline{r}, r^*)$, in which case $[\theta^*, \hat{\theta}) \subseteq \Theta(r)$, take any μ such that $\mu(\hat{\theta}|x,r) > g(r)$ if and only if $x < X(r)$. If $x < X(r)$, we also restrict $\mu(\{\theta \in \Theta(x) \cap \Theta(r)\}|x,r) = 0$ when $\Theta(x) \nsubseteq \Theta(r)$. If instead $x \geq X(r)$, we do not need to impose such a restriction, for it would only make the agents more aggressive and hence the deviation even less profitable. Finally, when either $r > r^*$, or $r = r^*$ and $\Theta(x) \cap \Theta^{**} = \varnothing$, necessarily $\Theta(x) \subseteq \Theta(r)$. Take then any beliefs such that $\mu(0|x,r) = 0$ for $x \geq \underline{x}$ and $\mu(0|x,r) = 1$ otherwise. Given these beliefs and the definition of X, the strategy of the agents is indeed sequentially rational. *QED*

References

[1] Albanesi, Stefania, V.V. Chari, and Lawrence Christiano (2003), "Expectation Traps and Monetary Policy," *Review of Economic Studies* 70, 715-741.

[2] Angeletos, George-Marios, Christian Hellwig, and Alessandro Pavan (2003) "Coordination and Policy Traps," NBER Working Paper 9767.

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

[4] Battaglini, Marco, and Roland Benabou (2003), "Trust, Coordination, and the Industrial Organization of Political Activism," *Journal of the European Economic Association* 1, 851-870.

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

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

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

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

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

[10] Cole, Harold, and Timothy Kehoe (2000), "Self-fulfilling Debt Crises," *Review of Economic Studies* 67, 91-116.

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

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

[13] Corsetti, Giancarlo, Bernardo Guimaraes and Nouriel Roubini, (2004), "International Lending of Last Resort and Moral Hazard: A Model of IMF's Catalytic Finance," working paper, University of Rome.

[14] Dasgupta, Amil (2003), "Coordination, Learning and Delay," working paper, London School of Economics.

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

[16] Drazen, Allen (2000), "Interest-Rate and Borrowing Defense Against Speculative Attack," *Carnegie-Rochester Conference Series on Public Policy* 52, 303-348.

[17] Goldstein, Itay, and Ady Pauzner (2000), "Demand Deposit Contracts and the Probability of bank Runs," working paper, Duke University and Tel Aviv University.

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

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

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

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

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

[23] Morris, Stephen, and Hyun Song Shin (2003), "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.

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

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

[26] Rochet, Jean-Charles, and Xavier Vives (2004), "Coordination Failures and the Lender of Last Resort: Was Bagehot Right After All?," working paper, University of Toulouse.

[27] Rogoff, Kenneth, and Anne Sibert (1988), "Elections and Macroeconomic Policy Cycles," *Review of Economic Studies* 55, 1-16.

Coordination and Policy Traps: Online Appendix

George-Marios Angeletos MIT and NBER

Christian Hellwig UCLA

Alessandro Pavan Northwestern University

A1. Negative correlation between the value and the strength of the status quo

Proposition 6 in the paper assumes that the payoff the policy maker enjoys from maintaining the status quo is positively correlated with (or independent of) its strength. The following example shows that such a positive correlation is not essential.

Proposition A1. *Suppose $R(\theta, A, r) = \theta - A$ and $U(\theta, A, r) = v(\theta) - A - C(r)$, where v is not necessarily monotonic, but satisfies $v(\theta) > 1 - \underline{r}$ for all $\theta \in [0, 1)$. There exists $\hat{r} > \underline{r}$ such that, for any $r^* \in [\underline{r}, \hat{r})$, there is an equilibrium in which the status quo is abandoned if and only if $\theta < 0$ and the policy maker sets r^* for $\theta \in [0, \theta^{**}]$ and $\underline{r}$ otherwise.*

Proof. Let $\hat{r} \in (\underline{r}, \tilde{r})$ be the unique solution to

$$C(\hat{r}) = \sigma \left[\Psi^{-1}\left(1 - \tfrac{\underline{r}}{1-\underline{r}} C(\hat{r})\right) - \Psi^{-1}(C(\hat{r}))\right]$$

and note that $C(\hat{r}) < 1 - \underline{r}$. For any $r^* \in (\underline{r}, \hat{r})$, let

$$\begin{aligned} x^* &= \sigma \Psi^{-1}\left(1 - \tfrac{\underline{r}}{1-\underline{r}} C(r^*)\right) \\ \theta^{**} &= x^* - \sigma \Psi^{-1}(C(r^*)) = \sigma \left[\Psi^{-1}\left(1 - \tfrac{\underline{r}}{1-\underline{r}} C(r^*)\right) - \Psi^{-1}(C(r^*))\right] \end{aligned} \tag{1}$$

and note that $\theta^{**} \geq 0$ for any $r^* < \hat{r}$ and solves $\Psi(\frac{x^* - \theta^{**}}{\sigma}) = C(r^*)$. Finally, let $\hat{\theta} \in (0, 1)$ be the unique solution to $\Psi(\frac{x^* - \hat{\theta}}{\sigma}) = \hat{\theta}$ and observe that $\theta^{**} > \hat{\theta}$ since $\Psi(\frac{x^* - \theta^{**}}{\sigma}) < \theta^{**}$ when $r^* < \hat{r}$.

We next prove that the following is an equilibrium: the policy maker sets $r(\theta) = r^*$ if $\theta \in [0, \theta^{**}]$ and $r(\theta) = \underline{r}$ otherwise; agents attack if and only if $(x, r) < (x^*, r^*)$, or $x < \underline{x}$; and the status quo is abandoned if and only if $\theta < 0$.

Consider first the behavior of the agents. For $r = \underline{r}$, beliefs are pinned down by Bayes' rule (this is immediate when noise is unbounded, whereas with bounded noise, it follows from the fact that $\theta^{**} < 2\sigma$) and satisfy $\mu\left(0|x,\underline{r}\right) > \underline{r}$ if and only if $x < x^*$, where x^* solves

$$\frac{1-\Psi(\frac{x^*}{\sigma})}{1-\Psi(\frac{x^*}{\sigma})+\Psi(\frac{x^*-\theta^{**}}{\sigma})} = \underline{r}. \tag{2}$$

For any (x,r) such that $r = r^*$ and $\Theta\left(x\right) \cap [0,\theta^{**}] \neq \varnothing$, μ is also determined by Bayes' rule and satisfies $\mu\left(0|x,\underline{r}\right) = 0$. For any (x,r) such that either $r = r^*$ and $\Theta\left(x\right) \cap [0,\theta^{**}] = \varnothing$, or $r > r^*$, $\Theta\left(x\right) \subseteq \Theta(r)$, take any beliefs such that $\mu\left(0|x,r^*\right) = 1$ if $x < \underline{x}$ and $\mu\left(0|x,r^*\right) = 0$ otherwise. Finally, for any $r \in (\underline{r}, r^*)$, note that $[0,\theta^{**}] \cap \Theta(r) = \varnothing$. Then take any beliefs such that $\mu(\hat{\theta}|x,r) > \underline{r}$ if and only if $x < x^*$ and $\mu(\{\theta \in \Theta(x) \cap \Theta(r)\}|x,r) = 0$ if $\Theta(x) \nsubseteq \Theta(r)$. Given these beliefs, the strategy of the agents is sequentially rational for any (x,r).

Consider next the policy maker. Given the strategy of the agents, it is optimal to set either $\underline{r}$ or r^*. The payoff from setting $\underline{r}$ is zero for $\theta \leq \hat{\theta}$ and $v\left(\theta\right) - \Psi(\frac{x^*-\theta}{\sigma})$ for $\theta > \hat{\theta}$, whereas the payoff from setting r^* is negative for $\theta < 0$ and $v\left(\theta\right) - C\left(r^*\right)$ for $\theta \geq 0$. Since $\Psi(\frac{x^*-\theta^{**}}{\sigma}) = C\left(r^*\right) \leq C(\hat{r}) < 1 - \underline{r} < v\left(\theta\right)$, it follows that r^* is optimal if and only if $\theta \in [0,\theta^{**}]$, which completes the proof. *QED*

The above result assumes that v is sufficiently high. Multiplicity, however, survives even if v is negative for all θ : there exists a continuum of equilibria in which an intermediate set of θ who would maintain the status quo even by setting $\underline{r}$, prefer to raise the policy at r^*, because the cost of the policy is lower than that of the attack at $\underline{r}$ (i.e., $C\left(r^*\right) \leq A\left(\theta,\underline{r}\right)$). These equilibria differ with respect to both the level of the policy and the regime outcome.

A2. Incomplete information vs common knowledge

In this section we analyze the variant of the benchmark model in which V satisfies $V_{\theta A} \geq 0$ and $\lim_{\theta\to\infty}[V\left(\theta,0\right) - V\left(\theta,1\right)] = 0$ and ψ is log-concave and strictly positive over $\mathbb{R}$. The purpose of the exercise here is to contrast the set of equilibrium outcomes sustained under incomplete information with that under common knowledge.

First, we prove the analogues of Propositions 2 and 3.

Proposition A2. **(i)** *There exists an inactive-policy equilibrium in which* $r\left(\theta\right) = \underline{r}$ *for all* θ, *agents attack if and only if* $x < \tilde{x}$, *and* $D\left(\theta\right) = 1$ *if and only if* $\theta < \tilde{\theta}$, *where* $\tilde{\theta}$ *solves* $V(\tilde{\theta}, 1-\underline{r}) = 0$ *and* $\tilde{x} = \tilde{\theta} + \sigma\Psi^{-1}\left(1-\underline{r}\right)$. **(ii)** *For any* $r^* \in (\underline{r}, \tilde{r}]$, *there exist unique* $\theta^* \in (0, \tilde{\theta}]$, $\theta^{**} \geq \theta^*$, *and* $x^* \in \mathbb{R}$, *and an active-policy equilibrium in which* $r(\theta) = r^*$ *if* $\theta \in [\theta^*, \theta^{**}]$ and $r(\theta) = \underline{r}$ *otherwise, agents attack if and only if* $x < \underline{x}$ *or* $(x, r) < (x^*, r^*)$, *and* $D(\theta) = 1$ *if and only if* $\theta < \theta^*$.

Proof. Part (i) follows from the same arguments as in Proposition 2. Thus consider part (ii). Fix an arbitrary $r^* > \underline{r}$ and let $\theta^* > 0$ be the unique solution to $V(\theta^*, 0) = C(r^*)$. Next, for any $\theta^{\circ\circ} \geq \theta^*$, let $x^\circ = X(\theta^*, \theta^{\circ\circ})$ be the unique solution to $m(x^\circ; \theta^*, \theta^{\circ\circ}) = \underline{r}$, where

$$m(x; \theta^*, \theta^{\circ\circ}) \equiv \frac{1 - \Psi(\frac{x-\theta^*}{\sigma})}{1 - \Psi(\frac{x-\theta^*}{\sigma}) + \Psi(\frac{x-\theta^{\circ\circ}}{\sigma})} = \left[1 + \frac{\Psi(\frac{x-\theta^{\circ\circ}}{\sigma})}{1 - \Psi(\frac{x-\theta^*}{\sigma})}\right]^{-1}$$

and let $B(\theta^*, \theta^{\circ\circ}) \equiv \Psi(\frac{X(\theta^*,\theta^{\circ\circ})-\theta^{\circ\circ}}{\sigma})$. $B(\theta^*, \theta^{\circ\circ})$ is decreasing in $\theta^{\circ\circ}$ with maximal value $B(\theta^*, \theta^*) = 1 - \underline{r}$. Next, let

$$G(\theta^*, \theta^{\circ\circ}) \equiv V(\theta^{\circ\circ}, 0) - V(\theta^{\circ\circ}, B(\theta^*, \theta^{\circ\circ})) - V(\theta^*, 0)$$

$G(\theta^*, \theta^{\circ\circ})$ is decreasing in $\theta^{\circ\circ}$ (by the assumptions that $V_A < 0 < V_\theta$, $V_{\theta A} \geq 0$, and the monotonicity of B), with $G(\theta^*, \theta^*) = -V(\theta^*, 1 - \underline{r})$ and $G(\theta^*, \infty) = -V(\theta^*, 0) < 0$ (by the limit condition). It follows that a solution to $G(\theta^*, \theta^{\circ\circ}) = 0$ is unique whenever it exists; and it exists if and only if $G(\theta^*, \theta^*) \geq 0$, or equivalently $\theta^* \leq \tilde{\theta}$ (i.e., $r^* \leq \tilde{r}$). Let then θ^{**} be this solution and $x^* = X(\theta^*, \theta^{**})$. With $(x^*, \theta^*, \theta^{**})$ defined as above, the rest of the proof follows from the same arguments as in Proposition 3. *QED*

Next, we show that the equilibria identified above exhaust the set of equilibrium outcomes. When no $r \neq \underline{r}$ is played in equilibrium, we have the pooling equilibrium; hence, in what follows, we consider equilibria in which $r(\theta) > \underline{r}$ for some θ.

Proposition A3. *In any equilibrium in which* $\{\theta : r(\theta) > \underline{r}\} \neq \varnothing$, *there is* $r^* \in (\underline{r}, \tilde{r}]$ *such that the following hold:* $r(\theta) = r^*$ *for all* $\theta \in [\theta^*, \theta^{**}]$ *and* $r(\theta) = \underline{r}$ otherwise; *agents do not attack when they observe* $r = r^*$ *and attack if and only if* $x < x^*$ *when they observe* $r = \underline{r}$; $D(\theta) = 1$ *if and only if* $\theta < \theta^*$; *the thresholds* $(x^*, \theta^*, \theta^{**})$ *are unique and defined as in Proposition A2.*

Proof. We prove this claim with a series of Lemmas.

Lemma A1. *There is at most one* $r^* \neq \underline{r}$ *such that* $r(\theta) = r^*$ *whenever* $r(\theta) \neq \underline{r}$.

Proof. Since raising the policy and abandoning the regime is strictly dominated for all θ, any $r > \underline{r}$ that is played in equilibrium by some θ must lead to no regime change for this θ. But since the noise is unbounded, any θ can ensure no attack by playing such an r. And since C is strictly increasing, there can be at most one such r played in equilibrium. □

We henceforth fix some $r^* > \underline{r}$ and consider the set of equilibria in which r^* is played. Given such an equilibrium, let $I(\theta)$ be an indicator of whether θ raises the policy (i.e., $I(\theta) = 0$ if

$r(\theta) = \underline{r}$ and $I(\theta) = 1$ if $r(\theta) = r^*$) and, provided $\{\theta : I(\theta) = 1\} \neq \varnothing$, let $\theta' = \inf\{\theta : I(\theta) = 1\}$ and $\theta'' = \sup\{\theta : I(\theta) = 1\}$.

Lemma A2. *An equilibrium with $r^* > \underline{r}$ exists only if $r^* \leq \tilde{r}$ and satisfies $\theta^* \leq \theta' \leq \theta'' \leq \theta^{**}$.*

Proof. Clearly, $I(\theta) = 0$ for any $\theta < \theta^*$. Moreover, $\lim_{\theta \to \infty} [V(\theta, 0) - V(\theta, 1)] = 0$ ensures that $V(\theta, 0) - V(\theta, 1) < C(r^*)$ and therefore $I(\theta) = 0$ for θ sufficiently high. Hence, whenever $\{\theta : I(\theta) = 1\} \neq \varnothing$, necessarily $\theta^* \leq \theta' \leq \theta'' < \infty$.

Since any $\theta > \theta^*$ can always set r^*, face no attack, and ensure a payoff $V(\theta, 0) - C(r^*) > 0$, necessarily $D(\theta) = 0$ for all $\theta > \theta^*$. Let $\delta(x)$ denote the probability conditional on x that $\theta < \theta^*$ and $D(\theta) = 0$ and $p(x)$ the probability conditional on x that $\theta \in [\theta^*, \theta'']$ and $r(\theta) = \underline{r}$. Then, the posterior probability of regime change conditional on x and $\underline{r}$ is

$$\frac{1 - \Psi(\frac{x-\theta^*}{\sigma}) - \delta(x)}{1 - \Psi(\frac{x-\theta^*}{\sigma}) + p(x) + \Psi(\frac{x-\theta''}{\sigma})} \leq \frac{1 - \Psi(\frac{x-\theta^*}{\sigma})}{1 - \Psi(\frac{x-\theta^*}{\sigma}) + \Psi(\frac{x-\theta''}{\sigma})} \equiv m(x; \theta^*, \theta'').$$

Recalling the definition of m, X, and B from Claim 1, it follows that $x \geq X\left(\theta^*, \theta''\right)$ suffices for $\mu(D = 1 | x, \underline{r}) \leq \underline{r}$ and therefore $A(\theta, \underline{r}) \leq \Psi(\frac{X(\theta^*, \theta'') - \theta}{\sigma})$ for any θ.

If $\{\theta : I(\theta) = 1\} \neq \varnothing$, θ'' must be no less than θ^* and finite, and must solve the indifference condition $V\left(\theta'', 0\right) - C(r^*) = V\left(\theta'', A\left(\theta'', \underline{r}\right)\right)$. Since $A\left(\theta'', \underline{r}\right) \leq \Psi(\frac{X(\theta^*, \theta'') - \theta''}{\sigma}) = B\left(\theta^*, \theta''\right)$, this implies $V\left(\theta'', 0\right) - C(r^*) \geq V\left(\theta'', B\left(\theta^*, \theta''\right)\right)$, or equivalently $G\left(\theta^*, \theta''\right) \geq 0$. But if $r^* > \tilde{r}$, we know by the proof of Claim 1 that $G(\theta^*, \theta) < 0$ for all $\theta \geq \theta^*$, which gives a contradiction and proves that $\{\theta : I(\theta) = 1\} = \varnothing$. That is, there is no equilibrium in which $r^* > \tilde{r}$. If instead $r^* < \tilde{r}$, $G\left(\theta^*, \theta''\right) \geq 0$ together with the definition of θ^* implies $G\left(\theta^*, \theta''\right) \geq 0 = G(\theta^*, \theta^{**})$; and since G is decreasing in its second argument, we conclude that $\theta'' \leq \theta^{**}$. □

Lemma A3. $D(\theta) = 1$ *for all* $\theta < \theta^*$, $D(\theta) = 0$ *for all* $\theta > \theta^*$, *and* $\theta' = \theta^*$.

Proof. If r^* is played in equilibrium, it is necessary that agents do not attack whenever they observe $r = r^*$ and that their beliefs and strategies are such that the policy maker never finds it profitable to deviate to $r \notin \{\underline{r}, r^*\}$. Construct then a sequence $\{x_k, \theta_k\}_{k=0}^{\infty}$ as follows: let $\theta_0 = 0$; for $k \geq 0$, let x_k solve $1 - \Psi(\frac{x_k - \theta_k}{\sigma}) = \underline{r}$ and $\theta_{k+1} = \min\{\theta^*, \theta'_{k+1}\}$, where θ'_{k+1} solves $V(\theta'_{k+1}, \Psi(\frac{x_k - \theta'_{k+1}}{\sigma})) = 0$. This sequence is increasing and bounded from above; since V and Ψ are continuous, it converges to $(x_\infty, \theta_\infty)$, where $\theta_\infty = \min\{\theta^*, \tilde{\theta}\} = \theta^*$ (since $r^* \leq \tilde{r}$) and $x_\infty = \theta^* + \sigma \Psi^{-1}(1 - \underline{r})$. Note that $D(\theta) = 1$ for $\theta < \theta_0 (= 0)$, $D(\theta) = 0$ for $\theta \geq \theta^*$, and $I(\theta) = 0$ for $\theta < \theta^*$. Consider any $k \geq 0$ and suppose $D(\theta) = 1$ for all $\theta < \theta_k$. The posterior probability of regime change given signal x and policy $\underline{r}$ is then

$$\int_{\mathbb{R}} D(\theta)d\mu(\theta|x,\underline{r}) = \frac{\int_{-\infty}^{+\infty} D\left(\theta\right)\left[1-I\left(\theta\right)\right]\frac{1}{\sigma}\psi(\frac{x-\theta}{\sigma})d\theta}{\int_{-\infty}^{+\infty}\left[1-I\left(\theta\right)\right]\frac{1}{\sigma}\psi(\frac{x-\theta}{\sigma})d\theta} \geq 1-\Psi(\tfrac{x-\theta_k}{\sigma}),$$

implying that it is optimal for the agent to attack whenever $x < x_k$. But if agents attack whenever $x < x_k$, $A\left(\theta,\underline{r}\right) \geq \Psi(\frac{x_k-\theta}{\sigma})$ and therefore for any $\theta < \theta_{k+1}$, $D\left(\theta\right) = 1$ if θ sets $r = \underline{r}$. By induction then, $D\left(\theta\right) = 1$ for all $\theta < \theta^*$.

If $r^* = \tilde{r}$, then $\theta^* = \theta^{**} = \tilde{\theta}$ and therefore also $\theta' = \theta'' = \tilde{\theta}$. For any $r^* < \tilde{r}$, on the other hand, we have $\theta^* < \tilde{\theta}$, which together with the definition of $\tilde{\theta}$ and x_∞, gives $V(\theta^*, \Psi(\frac{x_\infty-\theta^*}{\sigma})) < V(\tilde{\theta}, \Psi(\frac{x_\infty-\theta^*}{\sigma})) = V(\tilde{\theta}, 1-\underline{r}) = 0$. By the continuity of V and Ψ then, there is $\delta > 0$ such that $V(\theta, \Psi(\frac{x_\infty-\theta}{\sigma})) < 0$ for all types $\theta \in [\theta^*, \theta^*+\delta)$; and since $V\left(\theta, 0\right) > C\left(r^*\right)$ for any $\theta > \theta^*$, these types necessarily set r^*, which proves that $\theta' = \theta^*$.□

So far we have established that, in any equilibrium in which r^* is played, $r\left(\theta\right) = r^*$ only if $\theta \in [\theta^*, \theta^{**}]$. It remains to show that $r\left(\theta\right) = r^*$ for *all* $\theta \in [\theta^*, \theta^{**}]$.

Lemma A4. *If $a(x,\underline{r})$ is decreasing in x, then $\theta'' = \theta^{**}$ and $r(\theta) = r^*$ for all $\theta \in [\theta^*, \theta^{**}]$.*

Proof. If $a(x,\underline{r})$ is decreasing in x, then necessarily $A(\theta,\underline{r})$ is decreasing in θ. But then $V\left(\theta, A\left(\theta,\underline{r}\right)\right) < V\left(\theta'', A\left(\theta'',\underline{r}\right)\right) = V\left(\theta'', 0\right) - C(r^*)$ for all $\theta < \theta''$ and therefore all $\theta \in [\theta^*, \theta'']$ necessarily set r^*. It follows that

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

and therefore an agent attacks if and only $x < X\left(\theta^*,\theta''\right)$, with m and X defined as in the proof of Claim 1. This in turn implies that $A(\theta,\underline{r}) = \Psi(\frac{X(\theta^*,\theta'')-\theta}{\sigma})$ and therefore θ'' must solve $V(\theta'', \Psi(\frac{X(\theta^*,\theta'')-\theta}{\sigma})) = V\left(\theta'', 0\right) - C(r^*)$, or equivalently $G\left(\theta^*,\theta''\right) = 0$. But, by the definition of θ^{**} and the monotonicity of G, $G\left(\theta^*,\theta''\right) = 0$ if and only if $\theta'' = \theta^{**}$. □

Lemma A5. *If ψ is log-concave, $a(x,\underline{r})$ is decreasing in x.*

Proof. The probability of regime change given x and $\underline{r}$ is $\mu(\theta \leq \theta^*|x,\underline{r}) = \left(1+1/M\left(x\right)\right)^{-1}$, where

$$M(x) \equiv \frac{1-\Psi(\frac{x-\theta^*}{\sigma})}{\int_{\theta^*}^{\infty}\left[1-I\left(\theta\right)\right]\frac{1}{\sigma}\psi(\frac{x-\theta}{\sigma})d\theta}.$$

It follows that $\mu(\theta \leq \theta^*|x,\underline{r})$ is decreasing in x if $d\ln M\left(x\right)/dx < 0$, or equivalently

$$\frac{\int_{-\infty}^{\theta^*}\frac{1}{\sigma^2}\psi'\left(\frac{x-\theta}{\sigma}\right)d\theta}{\int_{-\infty}^{\theta^*}\frac{1}{\sigma}\psi\left(\frac{x-\theta}{\sigma}\right)d\theta} - \frac{\int_{\theta^*}^{\infty}\left[1-I\left(\theta\right)\right]\frac{1}{\sigma^2}\psi'\left(\frac{x-\theta}{\sigma}\right)d\theta}{\int_{\theta^*}^{\infty}\left[1-I\left(\theta\right)\right]\frac{1}{\sigma}\psi\left(\frac{x-\theta}{\sigma}\right)d\theta} < 0.$$

Using the fact that $I(\theta) = 0$ for all $\theta \leq \theta^*$, the above is equivalent to

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

which holds if ψ'/ψ is decreasing (i.e., if ψ is log-concave). The monotonicity of $\mu(\theta \leq \theta^*|x, \underline{r})$ then implies monotonicity of $a(x, \underline{r})$. □

Combining Lemmas A1-A5 completes the proof. *QED*

Finally, contrast the above result with the equilibrium outcomes sustainable under common knowledge ($\sigma = 0$). Any policy $r(\theta)$ such that $C(r(\theta)) \leq V(\theta, 0)$ for $\theta \in [0, 1]$ and $r(\theta) = \underline{r}$ otherwise can be part of a subgame perfect equilibrium. Moreover, the regime outcome is indeterminate for any $\theta \in [0, 1)$.

We conclude that incomplete information reduces the set of equilibrium outcomes as compared to common knowledge. This may permit interesting predictions: no $\theta > \tilde{\theta}$ abandons the status quo and the range of policy intervention vanishes when $\sigma \to 0$.

A3. Noise in the policy maker's observation of θ

In this appendix, we consider the case where the policy maker has imperfect information about the fundamentals: in stage 1, the policy maker does not observe θ; instead, she receives a signal $y = \theta + \eta\varepsilon$, where $\eta > 0$ parametrizes the quality of her information and ε is bounded noise, with support $[-1, 1]$, absolutely continuous c.d.f. G, and p.d.f. g.

That policy inaction can be sustained as an equilibrium is straightforward. We next show that any of the active-policy equilibria of Proposition 3 (where $\eta = 0$) can be approximated by an equilibrium in the game with $\eta > 0$. Hence, not only multiplicity survives, but also the same type of equilibria pertain.

Proposition A5. *For any $r^* \in (\underline{r}, \tilde{r})$, there exists $\bar{\eta} > 0$ such that for all $\eta < \bar{\eta}$, there exist (y', y'', x', θ') and an equilibrium in which $r(y) = r^*$ if $y \in [y', y'']$ and $r(y) = \underline{r}$ otherwise, agents attack if and only if $x < \underline{x}$ or $(r, x) < (r', x')$, and regime change occurs if and only if $\theta < \theta'$. As $\eta \to 0$, $(y', y'', x', \theta') \to (\theta^*, \theta^{**}, x^*, \theta^*)$, where the latter are defined as in Proposition 3.*

Proof. Consider first the policy maker. Given the strategy of the agents, any $r \neq \{\underline{r}, r^*\}$ is clearly dominated by either $\underline{r}$ or r^*. Now, suppose that, conditional on y', regime change occurs with certainty if the policy maker sets $\underline{r}$, whereas conditional on y'', the probability of regime change when setting $\underline{r}$ is zero. This is the case if and only if $y' + \eta < \theta^{\#} \leq y'' - \eta$, where $\theta^{\#}$ is the solution to $\theta^{\#} = \Psi(\frac{x' - \theta^{\#}}{\sigma})$. Then define y' and y'' by the following:

$$
\begin{aligned}
y' &= C(r^*) && (3) \\
C(r^*) &= \textstyle\int_{y''-\eta}^{y''+\eta} \Psi(\frac{x'-\theta}{\sigma})\frac{1}{\eta}g(\frac{y''-\theta}{\eta})d\theta && (4)
\end{aligned}
$$

Note that $\mathbb{E}[A(\theta, \underline{r})|y] = \int_{y-\eta}^{y+\eta} \Psi(\frac{x'-\theta}{\sigma})\frac{1}{\eta}g(\frac{y-\theta}{\eta})d\theta$ is strictly decreasing in y for any $y \in \mathbb{R}$ if ξ has unbounded support; if ξ is bounded, $\mathbb{E}[A(\theta, \underline{r})|y]$ is strictly decreasing in y for $y \notin [x' - \sigma - \eta, x' + \sigma + \eta]$, $\mathbb{E}[A(\theta, \underline{r})|y] = 1$ for $y \leq x' - \sigma - \eta$, and $\mathbb{E}[A(\theta, \underline{r})|y] = 0$ for $y \geq x' + \sigma + \eta$. Since $0 < C(r^*) \leq C(\tilde{r}) = \tilde{\theta} < 1$, a solution to $\mathbb{E}[A(\theta, \underline{r})|y''] = C(r^*)$, or equivalently (4), exists and is unique irrespectively of whether ξ is bounded or unbounded. As long as $y' - \eta > 0$ and $y' + \eta < \theta^{\#} \leq y'' - \eta$, which – as we prove below – hold for η small enough, the proposed strategy for the policy maker is optimal. If ξ is bounded, suppose further $|y'' - y'| < 2\sigma$, which again we will hold for η small enough.

Next, consider the behavior of the agents. When $r = \underline{r}$, beliefs are pinned down by Bayes rule; since $y' + \eta < \theta^{\#} \leq y'' - \eta$, the posterior probability of regime change for an agent with private signal x is

$$\Pr(\theta \leq \theta^{\#}|x, \underline{r}) = \Pr(y \leq y'|x, \underline{r}) = \frac{P_1(x; y')}{P_1(x; y') + P_2(x; y'')} = \frac{1}{1 + P_2(x; y'')/P_1(x; y')},$$

where

$$
\begin{aligned}
P_1(x; y) &\equiv 1 - \Psi(\tfrac{x-(y-\eta)}{\sigma}) + \int_{y-\eta}^{y+\eta} G(\tfrac{y-\theta}{\eta})\tfrac{1}{\sigma}\psi\left(\tfrac{x-\theta}{\sigma}\right)d\theta = \int_{y-\eta}^{y+\eta} [1 - \Psi(\tfrac{x-\theta}{\sigma})]\tfrac{1}{\eta}g(\tfrac{y-\theta}{\eta})d\theta \\
P_2(x; y) &\equiv \int_{y-\eta}^{y+\eta} [1 - G(\tfrac{y-\theta}{\eta})]\tfrac{1}{\sigma}\psi\left(\tfrac{x-\theta}{\sigma}\right)d\theta + \Psi(\tfrac{x-(y+\eta)}{\sigma}) = \int_{y-\eta}^{y+\eta} \Psi(\tfrac{x-\theta}{\sigma})\tfrac{1}{\eta}g(\tfrac{y-\theta}{\eta})d\theta
\end{aligned}
$$

x' thus solves the indifference condition $\Pr(y \leq y'|x', \underline{r}) = \underline{r}$, or equivalently

$$P_1\left(x'; y'\right) = \tfrac{\underline{r}}{1-\underline{r}} P_2\left(x'; y''\right). \tag{5}$$

Note that $P_1(x; y')$ is decreasing in x, while $P_2(x; y'')$ is increasing in x, implying that $\Pr(y \leq y'|x, \underline{r})$ is decreasing in x. Moreover, $\Pr(y \leq y'|x, \underline{r}) \to 1$ as $x \to -\infty$ and $\Pr(y \leq y'|x, \underline{r}) \to 0$ as $x \to +\infty$, which ensures the existence of a unique solution to (5) given y' and y''.

Substituting $C(r^*) = P_2(x'; y'')$ from (4) into (5) and using (3) gives a single equation in x':

$$P_1\left(x'; C(r^*)\right) = \tfrac{\underline{r}}{1-\underline{r}} C(r^*).$$

Since $P_1(x; C(r^*))$ is decreasing in x, with $\lim_{x\to-\infty} P(x; C(r^*)) = 1$ and $\lim_{x\to+\infty} P(x; C(r^*)) = 0$, the above admits a unique solution x'. Moreover, by the definition of x^* in the benchmark model, $\frac{r}{1-\underline{r}} C(r^*) = 1 - \Psi(\frac{x^*-C(r^*)}{\sigma})$. From the continuity of P_1 in η, it is then immediate that $x' \to x^*$ as $\eta \to 0$, in which case $\theta^\# \to \hat{\theta}$, $\mathbb{E}[A(\theta, \underline{r})|y''] \to \Psi(\frac{x^*-y''}{\sigma})$, and $y'' \to \theta^{**}$. Along with $y' = C(r^*) = \theta^*$, this implies that there exists $\bar{\eta} > 0$ such that $\eta < \bar{\eta}$ suffices for a solution to (3), (4) and (5) to exist and satisfy $y' - \eta > 0$, $y' + \eta < \theta^\# < y'' - \eta$, and, when ξ is bounded, $|y'' - y'| < 2\sigma$.

Finally, for out-of-equilibrium events, take any beliefs that satisfy the following: for $r \in (\underline{r}, r^*)$, $\mu(\theta^\#|x, r) > r$ if and only if $x < x'$; for $r > r^*$ or, in the case of bounded noise, $r = r^*$ and $x \notin [y' - \sigma, y'' + \sigma]$, $\mu(0|x, r) = 1$ if $x < \underline{x}$ and $\mu(0|x, r) = 0$ otherwise. Given these beliefs the strategy of the agents is sequentially rational. *QED*