Information Manipulation, Coordination and Regime Change*

Chris Edmond†

First draft: October 2003. This draft: March 2008

Abstract

This paper studies endogenous information manipulation in games where a population can overthrow a regime if individuals coordinate. The benchmark game has a unique equilibrium and in this equilibrium propaganda is effective if signals are sufficiently precise. Despite playing against perfectly rational individuals, a regime is able to manipulate information in a way that exploits heterogeneity in individual beliefs so that at equilibrium its chances of surviving are higher than they otherwise would be. This result is robust to alternative payoffs where the regime cares only for survival and to a number of alternative information structures, including situations where individuals have access to high-quality private information that is entirely uncontaminated by the regime.

Keywords: global games, hidden actions, signal-jamming, propaganda, bias.

JEL classifications: C72, D82, D84.

In coordination games, the information available to agents is a crucial determinant of equilibrium outcomes. Typically, the truthfulness of that information is never in question. But when some players choose what information to release, that information is likely to be biased.[1] The optimal amount of bias used by senders of information will depend on the way that receivers typically filter their information. In many settings, receivers find it easy to infer bias from their signals and information manipulation has no effect on equilibrium beliefs. In a coordination setting, however, the way that an individual receiver filters information will also depend on the filtering behavior of others. In equilibrium, senders may be able to manipulate information so as to exploit heterogeneity in receivers' beliefs and thereby affect outcomes.

*This is an extensively revised version of a chapter from my UCLA dissertation. Previous versions of this paper circulated under the title 'Information and the limits to autocracy'.

†Department of Economics, Stern School of Business, New York University. Email: ⟨cedmond@stern.nyu.edu⟩.
I would particularly like to thank Andrew Atkeson, Catherine de Fontenay, and Christian Hellwig for their encouragement and many helpful discussions. I also thank Marios Angeletos, Costas Azariadis, Heski Bar-Isaac, Adam Brandenburger, Ignacio Esponda, Stephen Morris, Alessandro Pavan, Andrea Prat, Hyun Shin, Adam Szeidl, Laura Veldkamp, Iván Werning and seminar participants at the ANU, Boston University, the University of Chicago, the Federal Reserve Bank of Minneapolis, La Trobe University, the London School of Economics, the University of Melbourne, MIT, Northwestern University, Oxford University, the University of Rochester, UC Irvine, University College London and UCLA for their comments and suggestions.

[1] Important recent contributions to the study of endogenous information in coordination games include Angeletos, Hellwig, and Pavan (2006, 2007) and Angeletos and Werning (2006).

This paper studies the use of biased information to manipulate outcomes in coordination games of regime change.[2] In these games, there are two outcomes of interest — either the preservation of a status quo or its overthrow.[3] I show that a regime with an interest in ensuring society coordinates on the status quo may be able to manipulate information to achieve that outcome in equilibrium.

Section 1 outlines the model. Individuals with an interest in preserving the status quo are treated as a single agent playing against a heterogeneous population of small agents. Each small player can take either an action that adversely affects the regime or an action that supports it. The actions of the small players are *strategic complements*. If enough people take the action that undermines the regime, then it is overturned. As in Morris and Shin (1998) and similar models, the ability of the regime to prevail is determined by a single parameter. Individuals are imperfectly informed about this parameter and in principle may coordinate either on overturning the regime or on living with the status quo. In this paper, however, the regime is endowed with the ability to take a costly *hidden action* that influences the distribution of information. In the benchmark model, this *signal-jamming* technology allows the regime to shift the mean of the distribution from which individuals sample so that individuals receive information that at face-value suggests the regime is difficult to overthrow. Rational individuals understand that their information is contaminated by the regime's propaganda and evaluate their signals accordingly.

This paper makes two contributions. I first show that this coordination game with endogenous information manipulation has a unique equilibrium. I then show that signal-jamming is effective in equilibrium when the intrinsic *precision* of the signal distribution is high. As signal precision becomes increasingly high, so does the ex ante probability of the regime surviving. This suggests that a regime's propaganda apparatus will be more useful to it when individuals are receiving, from a technological standpoint, intrinsically high quality signals. Perhaps the information technology revolution may not be as threatening to autocratic regimes as is sometimes supposed.[4]

Equilibrium uniqueness is proved in Section 2. A key intermediate result is that, despite the regime's manipulation, in any candidate equilibrium individuals' assessments of the probability of the regime being overthrown are monotonic in their signals. This makes it possible to use 'global games' arguments along the lines of Carlsson and van Damme (1993) and Morris and Shin (1998). In particular, it is possible to first show that there exists a unique equilibrium when strategies are monotone and then to show that this is the only equilibrium which survives the iterative elimination of strictly dominated strategies.

The effectiveness of signal-jamming when signals are precise is proved in Section 3. In equilib-

[2] Cheli and Della-Posta (2002) study a coordination game where individuals' signals are exogenously biased. By contrast, this paper studies an equilibrium problem where the amount of bias introduced by the regime has to be compatible with rational beliefs on the part of information receivers.

[3] The underlying coordination game employed in this paper has been used for many applications, including currency and debt crises, bank runs, and political unrest. See, for example, Obstfeld (1986, 1996), Morris and Shin (1998, 2003, 2004), Chamley (1999) and Atkeson (2000). See also Cooper (1999) and Chamley (2004, 237-267) for textbook overviews with many additional applications.

[4] See Kalathil and Boas (2003) for a discussion of the conventional wisdom on the role of information in undermining authoritarian regimes. They argue that the revolution in communication technologies like the internet is not likely to help in overthrowing such regimes.

rium, regimes are overthrown if their type is below an endogenous threshold. If a regime manipulates it generates a signal distribution with an artificially high mean that is strictly greater than the threshold. So if signals are precise, in this situation many individuals have signals suggesting the regime will survive. And consequently it is rational for any individual, when contemplating the beliefs of others, to assign relatively high probability to the event that they mostly have signals near this artificially high mean. At the margin this makes any individual less likely to attack and so the aggregate mass who do is relatively low. This in turn makes it more likely that the regime does in fact manipulate and create an artificially high signal mean thereby validating the original beliefs. Two things are critical to this argument. First, different types of regimes must take different actions so that there is uncertainty about the amount by which individuals should discount their signals: if all regimes took the same action, it would be easy to undo. Second, individuals must be imperfectly coordinated.

The equilibrium uniqueness result contrasts with Angeletos, Hellwig, and Pavan (2006), who were the first to emphasize endogenous information in global games. Angeletos, Hellwig, and Pavan (2006) have in mind situations where a regime's policy choice (e.g., a central bank's interest rate in anticipation of a speculative attack) provides an additional signal about the regime's type. Since individual strategies may or may not condition on the signal, this gives rise to both pooling and separating equilibria. In this paper, however, individuals' information is contaminated by a behind-the-scenes choice of the regime (e.g., pressure on a recalcitrant editor or general) that is commonly known to be possible but cannot be observed. Individuals have to disentangle the truth from the effects of the hidden action. In the benchmark model of Angeletos, Hellwig, and Pavan (2006) a pooling equilibrium occurs if all individuals believe that the regime will not intervene, for then the regime has no incentive to make a costly intervention. The regime is 'trapped' in an inactive policy equilibrium where the regime merely validates the belief that it will not intervene. But in this paper if all individuals believe that the regime will not take any action, the regime has a powerful incentive to manipulate information so there can be no pooling equilibrium.[5]

Traditional theories of strategic information transmission, such as Crawford and Sobel (1982) and Matthews and Mirman (1983), focus on a sole information receiver. In this paper there is instead a large cross-section of signal receivers. Since the incentives of the sender to manipulate depend on the aggregate behavior of the receivers and since individual actions are strategic complements, implicitly, each individual solves a filtering problem that depends simultaneously on the solutions to the filtering problems of all other receivers.

Section 4 shows that results are robust to alternative payoffs for the regime that give it no direct incentive to manipulate. Section 5 extends the model to cover alternative information structures, including situations where some information is *uncontaminated* by the regime's manipulation and an alternate technology for the regime where its actions directly affect signal precision.

[5]The relationship between this paper and an extension of Angeletos, Hellwig, and Pavan (2006) where individuals observe the regime's policy with idiosyncratic noise is discussed in Section 2 below.

1 Model of information manipulation and regime change

There is a unit mass of citizens, indexed by $i \in [0, 1]$. Citizens are ex ante identical. After drawing a signal (discussed below) each citizen decides whether to subvert the regime, $s_i = 1$, or not, $s_i = 0$. The population mass of subversives is $S := \int_0^1 s_i di$. If a citizen subverts, she pays a fixed opportunity cost $p > 0$.

The citizens face a regime indexed by a hidden state variable, denoted θ, that is the regime's private information. The type θ is normalized such that the regime is overthrown if and only if $\theta < S$. The payoff to a citizen is

$$u(s_i, S, \theta) = s_i(\mathbb{1}\{\theta < S\} - p) \tag{1}$$

where $\mathbb{1}$ denotes the indicator function. Individual actions s_i and the population aggregate S are *strategic complements*: the more citizens subvert the regime, the more likely it is that the regime is overthrown and so the more likely it is that any individual citizen's best action is to also subvert.

After learning θ, a regime may take a hidden action $a \geq 0$ that incurs a convex cost $C(a)$ where $C(0) = 0$, $C'(a) > 0$ for $a > 0$ and $C''(a) \geq 0$ for all a. I begin by assuming that the regime obtains a benefit $\theta - S$ from remaining in power. The regime is not just concerned with remaining in power but also wants to keep S small when it does survive. If $\theta < S$, the regime is overthrown and obtains an outside option with value normalized to zero. The payoff to a regime is therefore

$$B(S, \theta) - C(a) \tag{2}$$

where $B(S, \theta) := \mathbb{1}\{\theta \geq S\}(\theta - S)$.

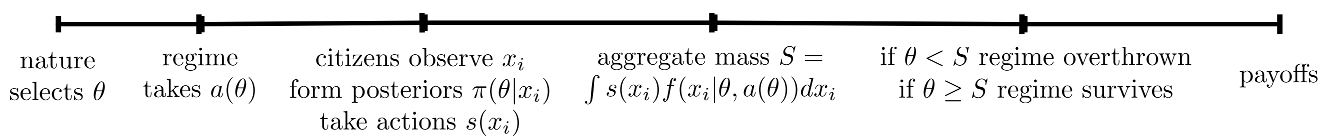

Figure 1: The timing of the model.

Following a regime's hidden action a, each citizen simultaneously draws an idiosyncratic signal $x_i := \theta + a + \varepsilon_i$ where the noise ε_i is independent of θ and is IID normally distributed with mean zero and *precision* α (that is, variance α^{-1}). So the density of signals is $f(x_i|\theta, a) := \sqrt{\alpha}\phi(\sqrt{\alpha}(x_i - \theta - a))$ where ϕ denotes the standard normal density. I begin by assuming that citizens have common priors for θ and that this prior is the (improper) uniform distribution over the whole real line.[6] The realization of the signal x_i is informative for both the type of the regime θ and the hidden action a. This action is itself informative about the regime's type and rational citizens take this into account when forming their beliefs. In equilibrium, the action taken by a regime and the beliefs of citizens will need to be mutually consistent. The timing of the model is shown in Figure 1.

[6] I discuss informative priors and heterogeneous priors in Section 5 below.

1.1 Equilibrium

A symmetric perfect Bayesian equilibrium is an individual's posterior density $\pi(\theta|x_i)$, individual subversion decision $s(x_i)$, mass of subversives $S(\theta, a)$ and hidden actions $a(\theta)$ such that

$$\pi(\theta|x_i) = \frac{f(x_i|\theta, a(\theta))}{\int_{-\infty}^{\infty} f(x_i|\theta, a(\theta))d\theta} \tag{3}$$

$$s(x_i) \in \underset{s_i\in\{0,1\}}{\operatorname{argmax}} \left\{ \int_{-\infty}^{\infty} u(s_i, S(\theta, a(\theta)), \theta)\pi(\theta|x_i)d\theta \right\} \tag{4}$$

$$S(\theta, a) = \int_{-\infty}^{\infty} s(x_i)f(x_i|\theta, a)dx_i \tag{5}$$

$$a(\theta) \in \underset{a\geq 0}{\operatorname{argmax}} \{B(S(\theta, a), \theta) - C(a)\} \tag{6}$$

The first condition says that a citizen with information x_i takes into account the regime's manipulation $a(\theta)$. The second says that given these beliefs, $s(x_i)$ is chosen to maximize its expected payoff. The third condition aggregates individual decisions to give the mass of subversives. The final condition says that the actions $a(\theta)$ maximize the regime's payoff. In equilibrium, the regime is overthrown if $R(\theta) := \mathbb{1}\{\theta < S(\theta, a(\theta))\} = 1$ while the regime survives if $R(\theta) = 0$.

1.2 Discussion of model

Regime's payoffs. Conditional on surviving a regime obtains benefit $\theta - S$ and so has a *direct* aversion to S. Regimes prefer to avoid a Prague Spring or a Tiananmen Square. There are two distinct motivations for this. First, suppressing a riot is likely to be resource costly to the regime and it seems natural for this cost to be increasing in the mass of rioters. Second, loosely speaking, a benefit $\theta - S$ provides a simple way of modeling aversion to information generated by observations of large S. For example, in a dynamic version of this model citizens would observe (possibly with noise) the mass of people who have in the past attempted to overthrow the regime. The size of previous attacks contains information about θ — in particular, low S is associated with high θ — and so by keeping S low even when it will survive this period, the regime may have another tool for convincing individuals not to attack.[7] Section 4 shows that the main results of this paper go through when payoffs are instead determined by $B(S, \theta) = \mathbb{1}\{\theta \geq S\}v(\theta)$ so that the regime cares only about survival and attaches (gross) value $v(\theta)$ to survival. Two specifications are considered: (i) $v(\theta) = \theta$, and (ii) $v(\theta) = \overline{v} > 0$ all θ. In the former case the value of survival and the regime's strength, as measured by θ, are perfectly correlated. In the latter case, the value of survival and the regime's strength are independent.

Absence of free-rider problems. If a citizen is almost sure the regime will be overthrown she will subvert and will not *free-ride* on others. Although this seems unappealing, a bad free-rider problem is simply one of many factors that might lead to a high opportunity cost p. To see this formally,

[7] More crudely, observations of large S might be able to convince foreign powers that it would be easy to assist the regime's opponents it in bringing the regime down. The $\theta - S$ specification also facilitates comparison with Angeletos, Hellwig, and Pavan (2006), ensuring that the information structure is the only difference.

suppose citizens get a larger expected payoff if they were actively involved in the downfall of the regime.[8] Specifically, if the regime is overthrown a citizen gets random payoff $w \in \{\underline{w}, \overline{w}\}$ with $\underline{w} < \overline{w}$ and $\Pr(w = \underline{w}) := \nu(s_i)$ with $0 \leq \nu(1) < \nu(0) \leq 1$. That is, a citizen may free-ride but if they do so they face a higher probability of being caught and penalized by getting only $\underline{w}$. Then if the utility cost of subverting the regime is normalized to 1 and a citizen believes the regime will be overthrown with probability $P(x_i)$, she will participate if and only if

$$P(x_i)[\overline{w} - \nu(1)(\overline{w} - \underline{w})] - 1 \geq P(x_i)[\overline{w} - \nu(0)(\overline{w} - \underline{w})]$$

equivalently, if and only if

$$P(x_i) \geq \frac{1}{(\overline{w} - \underline{w})(\nu(0) - \nu(1))} =: p \tag{7}$$

If $p \geq 1$ it can never be rational for a citizen to participate in subversion. To make the model interesting, then, we need:

Assumption 1. The opportunity cost of subversion is not too high: $p < 1$.

A citizen with $P(x_i) < p$ chooses not to subvert, in part, because of the incentive to free-ride. Intuitively, this incentive is weak if the probability of being caught out is sufficiently high or if the penalty from being caught is sufficiently severe.

Interpretation of hidden actions. The hidden action $a \geq 0$ of the regime gives it the ability to choose the common component of individual beliefs and the potential to bias the information that citizens receive. If $a > 0$ citizens draw from a signal distribution that at face value suggests the regime will be more difficult to overthrow. This represents a situation where it is common knowledge that the regime is able to exert pressure on editors, force recalcitrant generals to stand on parade, etc — so as to depict itself as difficult to overthrow — but where it is not possible to observe that pressure or manipulation directly and it instead must be inferred. In Section 5 below I endow citizens with informative priors for θ, thereby giving individuals extra information that is independent of a.

1.3 Exogenous information benchmarks

Two important special cases of the model are when: (i) the regime's type is common knowledge, or (ii) hidden actions are prohibitively expensive. In each case, citizens have *exogenous information.*

If θ is common knowledge, costly hidden actions are pointless and $a(\theta) = 0$ all θ. The model reduces to the coordination game used by Obstfeld (1986, 1996) to discuss self-fulfilling currency crises. If $\theta < 0$, any crowd $S \geq 0$ can overthrow the regime. It is optimal for any individual to subvert, all do so, and the regime is overthrown. If $\theta \geq 1$, no crowd can overthrow the regime. It

[8] Perhaps it is more likely that a citizen will secure an influential position in the new regime if she participated in the overthrow of the old regime. Or perhaps retribution is exacted on those who are thought to have let others take the risks in overthrowing the regime. See, respectively, Jackson (2001) and Frommer (2005) for discussion of the retribution exacted on collaborators after the liberation of France and Czechoslovakia from Nazi rule.

is optimal for any individual to not riot, none do, and the regime survives. If $\theta \in [0,1)$, the regime is 'fragile' and multiple self-fulfilling equilibria can be sustained. For example, if each individual believes that everyone else will riot, it will be optimal for each citizen to do so and $S = 1 > \theta$ leads to the regime's overthrow and the vindication of the initial expectations.

If hidden actions are prohibitively expensive, $a(\theta) = 0$ all θ and each citizen has private signal $x_i = \theta + \varepsilon_i$. Because each citizen has a signal of the regime's type, expectations are no longer arbitrary. As discussed by Carlsson and van Damme (1993), Morris and Shin (1998) and subsequent literature, this introduces the possibility of pinning down a unique equilibrium outcome.[9] In this equilibrium, strategies are threshold rules: there is a unique type θ^* such that the regime is overthrown for $\theta < \theta^*$ and a unique signal x^* such that a citizen subverts for $x < x^*$.

Proposition 1. (Morris-Shin): The unique equilibrium thresholds $x^*_{\text{MS}}, \theta^*_{\text{MS}}$ simultaneously solve

$$\Phi\left[\sqrt{\alpha}\,(\theta^*_{\text{MS}} - x^*_{\text{MS}})\right] = p \tag{8}$$

$$\Phi\left[\sqrt{\alpha}\,(x^*_{\text{MS}} - \theta^*_{\text{MS}})\right] = \theta^*_{\text{MS}} \tag{9}$$

where Φ denotes the standard normal CDF. In particular, $\theta^*_{\text{MS}} = 1 - p$ independent of α.

The first condition says that if the regime's threshold is θ^*_{MS}, a citizen with signal $x_i = x^*_{\text{MS}}$ will be indifferent between subverting or not. The second condition says that if the signal threshold is x^*_{MS}, a regime with type $\theta = \theta^*_{\text{MS}}$ will be indifferent. In the analysis below, I will say that a regime's hidden action technology is *effective* if in equilibrium $\theta^* < \theta^*_{\text{MS}} = 1 - p$.

To understand the consequences of information sets that are a function of the regime's manipulation, we need to study a more difficult equilibrium problem where the regime's manipulation is not trivial and citizens internalize a regime's incentives.

2 Unique equilibrium with hidden actions

The first main result of this paper is:

Theorem 1. There is a unique perfect Bayesian equilibrium. The equilibrium is *monotone* in the sense that there exist thresholds x^* and θ^* such that $s(x_i) = 1$ for $x_i < x^*$ and zero otherwise while $R(\theta) = 1$ for $\theta < \theta^*$ and zero otherwise.

Details of the proof are given in Appendix A. The proof shows that (i) there is a unique equilibrium in monotone strategies, and (ii) that the unique monotone equilibrium is the only equilibrium which survives the iterative elimination of interim strictly dominated strategies.

2.1 Unique equilibrium in monotone strategies

Let $\hat{x} \in \mathbb{R}$ denote a candidate for the citizens' threshold in a monotone equilibrium and let $\theta_{\hat{x}}$ and $a_{\hat{x}}(\theta)$ denote candidates for the regime's threshold and hidden actions given $\hat{x}$.

[9] This result depends on a relatively diffuse common prior. See Hellwig (2002) and Morris and Shin (2000, 2003, 2004) for discussion of the possibility of multiple equilibria in coordination games when public information is sufficiently informative. Public information is introduced into the model with hidden actions in Section 5 below.

Regime's problem. Taking $\hat{x}$ as given the mass of citizens facing the regime is

$$S_{\hat{x}}(\theta + a) := \int_{-\infty}^{\hat{x}} f(x_i|\theta, a)dx_i = \Phi[\sqrt{\alpha}(\hat{x} - (\theta + a))] \tag{10}$$

Since the regime has access to an outside option normalized to zero, its problem can be written

$$V_{\hat{x}}(\theta) := \max[0, W_{\hat{x}}(\theta)] \tag{11}$$

where $W_{\hat{x}}(\theta)$ is the best payoff regime θ can get if it is not overthrown

$$W_{\hat{x}}(\theta) := \max_{a \geq 0} [\theta - S_{\hat{x}}(\theta + a) - C(a)] \tag{12}$$

LEMMA 1. For each $\hat{x} \in \mathbb{R}$, the unique solution to the regime's decision problem is characterized by (i) a threshold $\theta_{\hat{x}} \in [0, 1)$ such that the regime is overthrown if and only if $\theta < \theta_{\hat{x}}$ and (ii) a single-valued hidden action function $a_{\hat{x}} : \mathbb{R} \to \mathbb{R}_+$ with $a_{\hat{x}}(\theta) = 0$ for all $\theta < \theta_{\hat{x}}$.

Using the envelope theorem and the definition of $S_{\hat{x}}(\theta+a)$ in equation (10) shows that $W'_{\hat{x}}(\theta) > 1$ all θ. And since $W_{\hat{x}}(\theta) < 0$ for $\theta < 0$ and $W_{\hat{x}}(1) > 0$, by the intermediate value theorem there is a unique $\theta_{\hat{x}} \in [0, 1)$ such that $W_{\hat{x}}(\theta_{\hat{x}}) = 0$. Using (11), the regime is overthrown if and only if $\theta < \theta_{\hat{x}}$. Since positive actions are costly, the regime takes no action for $\theta < \theta_{\hat{x}}$. Otherwise, for $\theta \geq \theta_{\hat{x}}$, the actions of the regime solve

$$a_{\hat{x}}(\theta) \in \operatorname*{argmax}_{a \geq 0} [\theta - S_{\hat{x}}(\theta + a) - C(a)], \quad \theta \geq \theta_{\hat{x}} \tag{13}$$

and the threshold $\theta_{\hat{x}}$ is found from the indifference condition $W_{\hat{x}}(\theta_{\hat{x}}) = 0$, or more explicitly

$$\theta_{\hat{x}} = S_{\hat{x}}[\theta_{\hat{x}} + a_{\hat{x}}(\theta_{\hat{x}})] + C[a_{\hat{x}}(\theta_{\hat{x}})] \tag{14}$$

Taking $\hat{x}$ as given, (13)-(14) simultaneously determine the threshold $\theta_{\hat{x}}$ and the hidden actions $a_{\hat{x}}(\theta)$ that characterize the solution to the regime's problem.

Because of the additive signal structure a unit increase in θ and a unit increase in $\hat{x}$ perfectly offset each other in terms of their effect on the regime's desired action. Given this, hidden actions will depend only on the difference $\theta - \hat{x}$. This implies that the family of hidden action functions $a_{\hat{x}}(\theta)$ can be represented by a single function $a : \mathbb{R} \to \mathbb{R}_+$ that takes $\theta - \hat{x}$ as an argument. Although it involves a slight abuse of notation, it's convenient to acknowledge this alternative representation by writing $a_{\hat{x}}(\theta) = a(\theta - \hat{x})$.

Citizens' problem. Let $\hat{\theta}$ denote a candidate for the regime's threshold and let $\hat{a} : \mathbb{R} \to \mathbb{R}_+$ denote a candidate hidden action profile satisfying $\hat{a}(\theta) = 0$ for $\theta < \hat{\theta}$. Let $\hat{P}(\hat{\theta}, x_i)$ denote the probability assigned by a citizen with signal x_i to $\theta < \hat{\theta}$ when the hidden actions are $\hat{a}(\theta)$, so

$$\hat{P}(\hat{\theta}, x_i) := \Pr[\theta < \hat{\theta} \mid x_i,\ \hat{a}(\cdot)] = \frac{\int_{-\infty}^{\hat{\theta}} \sqrt{\alpha}\phi[\sqrt{\alpha}(x_i - \theta)]d\theta}{\int_{-\infty}^{\infty} \sqrt{\alpha}\phi[\sqrt{\alpha}(x_i - \theta - \hat{a}(\theta))]d\theta} \tag{15}$$

where the numerator uses $\hat{a}(\theta) = 0$ for $\theta < \hat{\theta}$. This probability has the properties:

LEMMA 2. For any $\hat{\theta} \in \mathbb{R}$ and any $\hat{a} : \mathbb{R} \to \mathbb{R}_+$ such that $\hat{a}(\theta) = 0$ for $\theta < \hat{\theta}$, the function $\hat{P} : \mathbb{R} \times \mathbb{R} \to [0,1]$ is continuous and:

(i) strictly increasing in $\hat{\theta}$ with $\hat{P}(-\infty, x_i) = 0$ and $\hat{P}(\infty, x_i) = 1$ for any x_i, and

(ii) strictly decreasing in x_i with $\hat{P}(\hat{\theta}, -\infty) = 1$ and $\hat{P}(\hat{\theta}, \infty) = 0$ for any $\hat{\theta}$.

Now instead of an arbitrary threshold $\hat{\theta}$ and hidden actions $\hat{a}(\theta)$, suppose almost all citizens subvert the regime when they have signals $x < \hat{x}$ and the corresponding solution to the regime's problem is given by threshold $\theta_{\hat{x}}$ and hidden actions $a_{\hat{x}}(\theta)$ from (13)-(14). For each $\hat{x} \in \mathbb{R}$ we can construct a posterior probability of the regime's overthrow analogous to (15). Write this probability as $P(\theta_{\hat{x}}, x_i, \hat{x})$ where

$$P(\theta_{\hat{x}}, x_i, \hat{x}) := \Pr[\theta < \theta_{\hat{x}} \mid x_i,\ a_{\hat{x}}(\cdot)] = \frac{\int_{-\infty}^{\theta_{\hat{x}}} \sqrt{\alpha}\phi[\sqrt{\alpha}(x_i - \theta)]d\theta}{\int_{-\infty}^{\infty} \sqrt{\alpha}\phi[\sqrt{\alpha}(x_i - \theta - a(\theta - \hat{x}))]d\theta} \tag{16}$$

and where the denominator uses the representation $a_{\hat{x}}(\theta) = a(\theta - \hat{x})$. Since $\theta_{\hat{x}}$ is pinned down by $\hat{x}$ through (13)-(14), this probability is a function only of $\hat{x}$ and the citizen's signal x_i. To acknowledge this, write $K(\hat{x}, x_i) := P(\theta_{\hat{x}}, x_i, \hat{x})$. Then:

LEMMA 3. The function $K : \mathbb{R} \times \mathbb{R} \to [0,1]$ is continuous and:

(i) strictly decreasing in x_i with $K(\hat{x}, -\infty) = 1$ and $K(\hat{x}, \infty) = 0$ for any $\hat{x}$, and

(ii) satisfies

$$K(\hat{x}, x_i) := P(\theta_{\hat{x}}, x_i, \hat{x}) = P(\theta_{\hat{x}} - \hat{x}, x_i - \hat{x}, 0) \tag{17}$$

for any $\hat{x} \in \mathbb{R}$. For $x_i = \hat{x}$ in particular $K(\hat{x}, \hat{x}) = P(\theta_{\hat{x}} - \hat{x}, 0, 0)$.

A citizen with signal x_i will subvert the regime if and only if $K(\hat{x}, x_i) \geq p$. Therefore given the solution to the regime's problem as implied by (13)-(14), the signal threshold $\hat{x}$ solves

$$K(\hat{x}, \hat{x}) = P(\theta_{\hat{x}} - \hat{x}, 0, 0) = p \tag{18}$$

Solving the problems simultaneously. Now define a function $\mu : \mathbb{R} \to [0,1]$ by $\mu(z) := P(z, 0, 0)$. This function does not depend on $\hat{x}$. And given Lemma 2 and we know μ is continuous, $\mu'(z) > 0$ all $z \in \mathbb{R}$ with $\mu(-\infty) = 0$ and $\mu(\infty) = 1$. Equation (18) therefore implies a unique value $\theta^* - x^*$ such that $\mu(\theta^* - x^*) = p$. Write this value

$$\theta^* - x^* = \mu^{-1}(p) \tag{19}$$

To show that there is a unique monotone equilibrium, we take this difference $\theta^* - x^*$ and plug it into the regime's indifference condition (14). Using $a_{\hat{x}}(\theta) = a(\theta - \hat{x})$ and the definition of the mass of subversives $S_{\hat{x}}(\theta + a)$, the regime's indifference condition (14) can be written

$$\theta_{\hat{x}} \quad = \quad \Phi[\sqrt{\alpha}(\hat{x} - \theta_{\hat{x}} - a(\theta_{\hat{x}} - \hat{x}))] + C[a(\theta_{\hat{x}} - \hat{x})] \tag{20}$$

The right hand side of (20) depends only on the difference $\theta_{\hat{x}} - \hat{x}$ and not on $\theta_{\hat{x}}$ or $\hat{x}$ independently. Because of this, we can plug $\theta^* - x^*$ into the right hand side of (20) to solve uniquely for the regime threshold $\theta^* \in [0, 1)$ so that we know both θ^* and the signal threshold $x^* \in \mathbb{R}$ separately. The equilibrium hidden actions are then $a(\theta - x^*)$.

2.2 Iterative elimination of strictly dominated strategies

We can now go on to show that there is no other equilibrium. The argument begins by showing that for sufficiently low signals it is a dominant strategy to subvert the regime and for sufficiently high signals it is a dominant strategy to not subvert the regime.

Dominance regions. If the regime has $\theta < 0$, any mass $S \geq 0$ can overthrow the regime. Similarly, if the regime has $\theta \geq 1$ it can never be overthrown. Any regime that is overthrown takes no action, since to do so would incur a cost for no gain. Similarly, any regime θ that is not overthrown takes an action no larger than the a such that $\theta = C(a)$. Any larger action must result in a negative payoff which can be improved upon by taking the outside option. Given this:

LEMMA 4. There exists a pair of signals $\underline{x} < \overline{x}$, both finite, such that $s(x_i) = 1$ is strictly dominant for $x_i < \underline{x}$ and $s(x_i) = 0$ is strictly dominant for $x_i > \overline{x}$.

The most *pessimistic* scenario for any citizen is that the regime is overthrown only if $\theta < 0$. Moreover in this scenario the largest hidden action that could be rational for regime θ is $\underline{a}(\theta) := C^{-1}(\theta)$ for $\theta \geq 0$ and zero otherwise. Given this, we can use Lemma 2 to show that there exist citizens with sufficiently low signals $x_i < \underline{x}$. For these citizens, irrespective of the actions or beliefs of other agents, it is (iteratively) strictly dominant to subvert the regime. Analogously, the most *optimistic* scenario for any citizen is that the regime is overthrown if $\theta < 1$. Moreover, the smallest hidden action that could be rational for regime θ in this scenario is $\overline{a}(\theta) := 0$. Again we can use Lemma 2 to show there exist citizens with sufficiently high signals $x_i > \overline{x}$ for whom it is strictly dominant to not subvert the regime.

Iterative elimination. Starting from the dominance regions implied by $\underline{x}, \overline{x}$ it is then possible to iteratively eliminate (interim) strictly dominated strategies. To begin this iteration we use:

LEMMA 5. If it is strictly dominant for $s(x_i) = 1$ for all $x_i < \underline{x}$, then the regime is overthrown for at least all $\theta < \theta_{\underline{x}}$ where $\theta_{\underline{x}} > 0$ solves

$$\theta_{\underline{x}} = S_{\underline{x}}[\theta_{\underline{x}} + a_{\underline{x}}(\theta_{\underline{x}})] + C[a_{\underline{x}}(\theta_{\underline{x}})], \qquad a_{\underline{x}}(\theta) \in \underset{a \geq 0}{\operatorname{argmax}} \left[\theta - S_{\underline{x}}(\theta + a) - C(a)\right], \quad \theta \geq \theta_{\underline{x}}$$

where $S_{\underline{x}}(z) = \Phi[\sqrt{\alpha}(\underline{x} - z)]$ for any $z \in \mathbb{R}$. Similarly, if it is strictly dominant for $s(x_i) = 0$ for all $x_i > \overline{x}$, the regime is not overthrown for at least all $\theta > \theta_{\overline{x}}$ where $\theta_{\overline{x}} < 1$ is defined analogously.

This provides the first step in iterating from $\underline{x}, \overline{x}$. Since the mass of citizens who subvert the regime is *at least* $S_{\underline{x}}(\theta + a)$ we can show that the situation facing citizens is less pessimistic than was assumed in calculating $\underline{x}$: the regime is overthrown for all $\theta < \theta_{\underline{x}}$ where $\theta_{\underline{x}} > 0$. Similarly, the mass of citizens who subvert the regime is *no more than* $S_{\overline{x}}(\theta + a)$ and this implies the situation is less optimistic than was assumed in calculating $\overline{x}$: the regime survives for all $\theta > \theta_{\overline{x}}$ where $\theta_{\overline{x}} < 1$.

The incentives for an individual citizen to subvert are at *their weakest* when they expect all other citizens to subvert only for $x_i < \underline{x}$. Considering the possibility that some citizens with $x_i \geq \underline{x}$ also subvert increases the regime threshold and the chances of the regime being overthrown. More specifically, since cumulative distribution functions are non-decreasing, for any beliefs of the citizens, the posterior probability assigned by a citizen with signal x_i to the regime's overthrow is at least as much as the probability they assign to $\theta < \theta_{\underline{x}}$. Equivalently, $K(\underline{x}, x_i) - p$ is *the most conservative* estimate of the expected gain to subverting. Lemma 3 and the intermediate value theorem imply there is a unique signal $\psi(\underline{x})$ solving

$$K[\underline{x}, \psi(\underline{x})] = p \tag{21}$$

such that if it is strictly dominant for $s(x_i) = 1$ for all $x_i < \underline{x}$, then it is strictly dominant for $s(x_i) = 1$ for *at least* all $x_i < \psi(\underline{x})$. A parallel argument establishes that if is it strictly dominant for $s(x_i) = 0$ for all $x_i > \overline{x}$, then it is strictly dominant for $s(x_i) = 0$ for at least all $x_i > \psi(\overline{x})$.

To see how we can use the function ψ to eliminate strictly dominated strategies, let $\underline{x}_1 = \psi(\underline{x})$. We can use Lemma 5 again to construct a new regime threshold[10] $\theta_{\underline{x}_1} > \theta_{\underline{x}}$ and, since a citizen with signal x_i assigns probability at least $K(\underline{x}_1, x_i)$ to the regime's overthrow, we can construct a new signal threshold $\psi(\underline{x}_1)$ solving $K[\underline{x}_1, \psi(\underline{x}_1)] = p$ such that if it is strictly dominant for $s(x_i) = 1$ for all $x_i < \underline{x}_1$, then it is strictly dominant for $s(x_i) = 1$ for at least all $x_i < \psi(\underline{x}_1)$.

The key remaining step to prove there is a unique perfect Bayesian equilibrium is to derive the fixed point properties of the function $\psi : \mathbb{R} \to \mathbb{R}$. These are given by:

Lemma 6. The function ψ is continuous and has a unique fixed point $x^* = \psi(x^*)$ with $\psi'(x^*) \in (0, 1)$. Moreover $\psi(x) \leq x^*$ for all $x < x^*$ and $\psi(x) \geq x^*$ for all $x > x^*$.

We build up a sequence $\{\underline{x}_n\}_{n=0}^{\infty}$ by iterating on ψ starting from the initial condition $\underline{x}_0 := \underline{x}$. Since ψ is continuous, has a unique fixed point $x^* = \psi(x^*)$ with $\psi'(x^*) \in (0, 1)$ and $\psi(\underline{x}_n) \leq x^*$ for all $\underline{x}_n < x^*$, this sequence is strictly monotone increasing, bounded above by x^* and converges to x^* in the limit. Similarly, the sequence $\{\overline{x}_n\}_{n=0}^{\infty}$ obtained by iterating on ψ starting from $\overline{x}_0 := \overline{x}$ is strictly monotone decreasing, bounded below by x^* and converges to x^* in the limit.

Intuitively, these sequences represent the elimination of strictly dominated strategies 'from below' and 'from above'. If no-one else subverts the regime, a citizen with signal $x_i < \underline{x}_1 = \psi(\underline{x}_0)$ will subvert the regime. Therefore *all* citizens with signals $x < \underline{x}_1$ have $s(x) = 1$. But if so, *at least* all citizens with signals less than $\underline{x}_2 = \psi(\underline{x}_1)$ also subvert the regime. And so on. After n iterations, the only candidates for a citizen's equilibrium strategy all have the property that $s(x_i) = 1$ for $x_i < \underline{x}_n$ and similarly $s(x_i) = 0$ for $x_i > \overline{x}_n$ with $s(x_i)$ arbitrary for $x_i \in [\underline{x}_n, \overline{x}_n]$. In the limit as $n \to \infty$, the only strategy that survives the elimination of strictly dominated strategies is the one

[10] Lemma 5 states that $\theta_{\underline{x}} > 0$ where 0 corresponds to the regime threshold in the most pessimistic scenario for the citizens. In using Lemma 5 starting instead with $\underline{x}_1$ it should be understood that the Lemma states $\theta_{\underline{x}_1} > \theta_{\underline{x}}$ where now $\theta_{\underline{x}}$ corresponds to the (iteratively) most pessimistic scenario.

with $s(x_i) = 1$ for $x_i < x^*$ and $s(x_i) = 0$ otherwise. In short, the unique monotone equilibrium is the only equilibrium.

2.3 Discussion

Signaling and signal-jamming. Theorem 1 contrasts with Angeletos, Hellwig, and Pavan (2006), who were the first to emphasize endogenous information in a global games context. In their benchmark model, individuals get one noisy observation of θ plus one observation of a signal a chosen at cost $C(a)$ by the regime which may also be informative for θ. Individual strategies $s(x_i, a)$ may condition on a. In this *signaling* game, there is typically an uninformative pooling equilibrium and many separating equilibria. For example, if each individual expects no manipulation, individual strategies and hence the aggregate mass S will be independent of a. Given this, the regime has no incentive to manipulate and so validates the original expectation.

The equilibrium multiplicity result in Angeletos, Hellwig, and Pavan (2006) is relatively robust in that they also provide examples where individuals receive one idiosyncratically noisy observation of θ and one idiosyncratically noisy signal of $a(\theta)$. They show that multiple equilibria arise even in this scenario where out-of-equilibrium beliefs play no role.[11] The essential difference between their model augmented with idiosyncratically noisy observations of the policy action $a(\theta)$ and this paper is that here individuals get one noisy observation of one object, the sum $\theta + a(\theta)$, instead of two separate signals of the two constituent parts.

Loosely speaking, this model is an example of a signal-jamming game. The presence of a signal-jamming technology raises the issue of whether the signal sender, the regime, is *effective* in deterring citizens from participating. This parallels the industrial organization literature on limit pricing and entry deterrence. In Milgrom and Roberts (1982) an incumbent firm may price below its static monopoly price to signal that it has low costs in the hope of deterring entry. This game has a pooling equilibrium and many separating equilibria where the incumbent's price is indeed below its monopoly price but where limit pricing is ineffective in that a rational entrant correctly infers the incumbent's costs from its price and so enters exactly when it would if the incumbent's costs were observable. By contrast, Matthews and Mirman (1983) study a signal-jamming problem where the entrant observes the incumbent's price with noise. They show there is often a unique equilibrium and that it exhibits both limit pricing and successful entry deterrence.[12]

Imperfectly coordinated signal receivers. In traditional models of strategic information transmission like Crawford and Sobel (1982) and Matthews and Mirman (1983) there is one sender and one receiver. But in this paper and in Angeletos, Hellwig, and Pavan (2006) there is a large cross-section of signal receivers. Since the incentives of the sender to manipulate depend on the aggregate behavior of the receivers and since individual actions are strategic complements, implicitly, each individual receiver's filtering problem depends simultaneously on the solution to the

[11] The action $a(\theta)$ also has a payoff relevant effect in Angeletos, Hellwig, and Pavan (2006) but this is not essential.
[12] See Wilson (1992), Riley (2001, 455-459) and Tirole (1988, 364-374) for further discussion.

filtering problem of all other receivers. If some citizens believe the regime is manipulating, they will conclude that the regime will survive (since only regimes with $\theta \geq \theta^*$ actively manipulate) and will not subvert. Because of the strategic complementarity, this will make other individuals less likely to subvert.

That receivers are imperfectly coordinated gives rise to effects absent from the traditional one sender/one receiver game. As discussed in Section 3, when precision is very high and citizens are imperfectly coordinated, signal-jamming is maximally effective: as the precision $\alpha \to \infty$, all the fragile regimes with $\theta \in [0, 1)$ survive. But if instead citizens are perfectly coordinated then all the regimes with $\theta \in [0, 1)$ are overthrown.

3 Equilibrium information manipulation

The second main result of this paper is that the regime's signal-jamming technology becomes more effective as the precision α increases. Section 3.1 prepares the ground by characterizing equilibrium hidden actions. Section 3.2 and Section 3.3 show that signal-jamming can be effective in equilibrium even when the regime is playing against rational citizens who internalize the regime's incentives. To give intuition for this, I first use a simple example with constant marginal costs before turning to more general convex costs.

Terminology. Throughout this section I draw a distinction between whether signal-jamming *occurs* in equilibrium (when $a(\theta) > 0$ for some θ) and whether it is *effective*. I measure the effectiveness of signal-jamming by its ability to reduce the regime's threshold θ^* relative to the Morris-Shin level of $\theta^*_{\text{MS}} = 1 - p$ (as in Proposition 1). A lower θ^* increases the regime's ex ante survival probability by making it more likely that nature draws a $\theta \geq \theta^*$. For high enough α the regime has $\theta^* < 1 - p$ so that endowing a regime with a signal-jamming technology allows it to survive in more states than it would in the Morris-Shin world.

I say that the regime *benefits* from lower θ^* even though this does not necessarily increase the regime's payoff. In principle, it might be the case that lower θ^* is achieved through large, costly, actions that give the regime a lower payoff than they would achieve in the Morris-Shin world. But it turns out that as precision $\alpha \to \infty$ and θ^* falls, hidden actions also become small so that the fall in θ^* represents a genuine increase in payoffs, at least in the limit.

3.1 Regime's hidden actions

In equilibrium, hidden actions $a(\theta)$ are characterized by the first order necessary condition[13]

$$C'(a) = \sqrt{\alpha}\phi[\sqrt{\alpha}(x^* - \theta - a)], \qquad \theta \geq \theta^* \tag{22}$$

The marginal benefit of an action is the associated reduction in the mass of subversives and at an interior solution this is equated to $C'(a)$. For signal-jamming to occur (meaning $a(\theta) > 0$ for at

[13]The first order condition (22) for $a(\theta) > 0$ may have zero, one or two solutions. In the event of two solutions, only the higher solution satisfies the second order condition.

least some θ), the cost function either has to be either (i) strictly convex, or (ii) if marginal costs are constant, $C'(a) = c$ all a, then the level of c cannot be 'too high': $c < \sqrt{\alpha}\phi(0) =: \overline{c}$. If either of these conditions is satisfied, then actions are zero for all $\theta < \theta^*$ before jumping up discontinuously to a positive value at the threshold θ^*. As the fundamentals of the regime become strong, costly actions taken to generate a favorable signal distribution begin to encounter diminishing returns and the action profile dies to zero. Figure 2 illustrates.

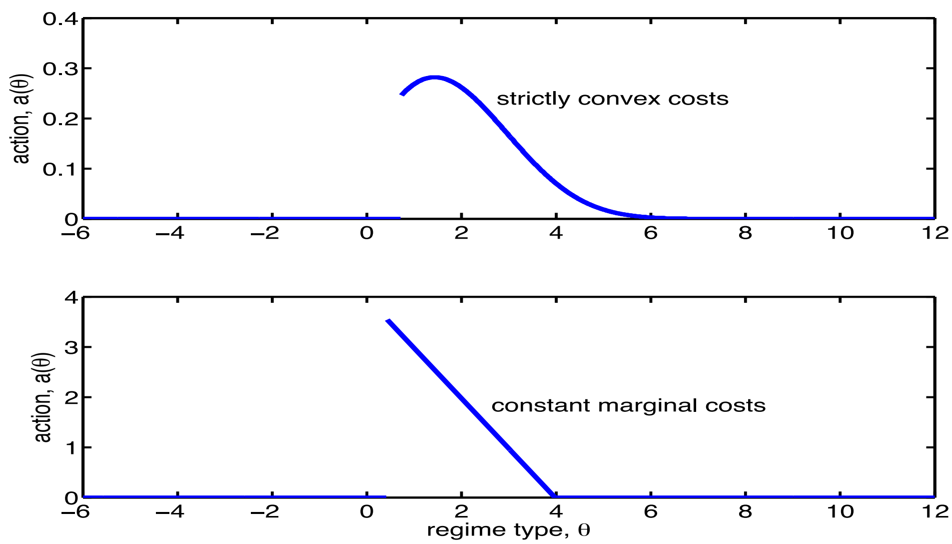

Figure 2: Equilibrium hidden actions $a(\theta)$.

Even if some regimes manipulate information in equilibrium, it is not necessarily the case that the existence of the signal-jamming technology is to the regime's benefit. To understand when a regime does benefit from the signal-jamming technology, I first specialize to the case of constant marginal costs. This example allows for explicit calculations showing that signal-jamming is effective when the signal precision α is sufficiently high.

3.2 Signal-jamming example: constant marginal costs

Let $C(a) := ca$ for some constant $c \in (0, \overline{c})$ where $\overline{c} := \sqrt{\alpha}\phi(0)$ so that $a(\theta) > 0$ for some θ. Then manipulating equation (22) shows that interior solutions to the regime's problem are given by

$$a(\theta) = x^* + \gamma - \theta, \qquad \theta \in [\theta^*, \theta^{**}) \tag{23}$$

where $\theta^{**} := x^* + \gamma$ and where

$$\gamma := \sqrt{\frac{2}{\alpha}\log\left(\frac{\sqrt{\alpha}\phi(0)}{c}\right)} > 0 \tag{24}$$

In this case, the signal-jamming is acute. All regimes that manipulate information pool on the same distribution of signals. Since the signal mean is $\theta + a(\theta)$, all regimes that manipulate, i.e.,

$\theta \in [\theta^*, \theta^{**})$, generate a mean of $\theta + x^* + \gamma - \theta = x^* + \gamma =: \theta^{**}$. These regimes mimic the signal mean of a regime type θ^{**} that is intrinsically more difficult to overthrow (than they are) and generate signals for the citizens $x_i = x^* + \gamma + \varepsilon_i$ that are *locally completely uninformative* about θ.

More precise information makes signal-jamming effective. In the absence of manipulation, higher precision α makes signals more tightly clustered around the true θ. Does this help citizens overthrow fragile regimes $\theta \in [0, 1)$ that survive only because of imperfect coordination? In the Morris-Shin case, the answer is 'no': the equilibrium threshold $\theta^*_{\text{MS}} = 1 - p$ is invariant to the precision of the private information.[14] With endogenous information manipulation, matters are even worse for the citizens. A regime's signal-jamming is more effective for high values of α than for low values.

To see this formally, use the definition of a citizen's posterior probability of overthrowing the regime (16) and write the indifference condition for the marginal citizen evaluated at x^*, θ^* as

$$\frac{\Phi[\sqrt{\alpha}(\theta^*_\alpha - x^*_\alpha)]}{\Phi[\sqrt{\alpha}(\theta^*_\alpha - x^*_\alpha)] + \int_{\theta^*_\alpha}^{\infty} \sqrt{\alpha}\phi[\sqrt{\alpha}(x^*_\alpha - \theta - a(\theta))]d\theta} = p$$

where a subscript α acknowledges dependence on the signal precision. Rearranging gives

$$\Phi[\sqrt{\alpha}(\theta^*_\alpha - x^*_\alpha)] = \frac{p}{1-p}\int_{\theta^*_\alpha}^{\infty} \sqrt{\alpha}\phi[\sqrt{\alpha}(x^*_\alpha - \theta - a(\theta))]d\theta \tag{25}$$

Now use the first order condition (22) and $C'(a) = c$ to simplify the right hand side integral

$$\begin{aligned}\int_{\theta^*_\alpha}^{\infty} \sqrt{\alpha}\phi[\sqrt{\alpha}(x^*_\alpha - \theta - a(\theta))]d\theta &= \int_{\theta^*_\alpha}^{\theta^{**}_\alpha} c\,d\theta + \int_{\theta^{**}_\alpha}^{\infty} \sqrt{\alpha}\phi[\sqrt{\alpha}(x^*_\alpha - \theta)]d\theta \\ &= c(x^*_\alpha - \theta^*_\alpha + \gamma_\alpha) + \Phi(-\sqrt{\alpha}\gamma_\alpha)\end{aligned}$$

where the first equality uses $a(\theta) = 0$ for $\theta \geq \theta^{**}_\alpha$ from (23) and the second equality uses $\theta^{**}_\alpha = x^*_\alpha + \gamma_\alpha$. Plugging this back into (25) gives us the first of two equations characterizing the two unknown thresholds

$$\Phi[\sqrt{\alpha}(\theta^*_\alpha - x^*_\alpha)] = \frac{p}{1-p}\left[c(x^*_\alpha - \theta^*_\alpha + \gamma_\alpha) + \Phi(-\sqrt{\alpha}\gamma_\alpha)\right] \tag{26}$$

The second equation comes from the regime's indifference condition (14) and can be written

$$\theta^*_\alpha = c(x^*_\alpha - \theta^*_\alpha + \gamma_\alpha) + \Phi(-\sqrt{\alpha}\gamma_\alpha) \tag{27}$$

For each $\alpha > 0$, the two equations (26)-(27) uniquely determine the two thresholds $x^*_\alpha, \theta^*_\alpha$ [solve (26) for the unique difference $\theta^*_\alpha - x^*_\alpha$ and then plug into (27) to get θ^*_α]. The equilibrium mass of subversives that makes the regime indifferent is $S^*_\alpha := \Phi(-\sqrt{\alpha}\gamma_\alpha) = \Phi[-\sqrt{2\log(\sqrt{\alpha}\phi(0)/c)}]$ which is strictly decreasing in α and $S^*_\alpha \to 0^+$ as $\alpha \to \infty$. High α helps the regime engineer a small mass

[14]See Proposition 1 for more discussion. The invariance of the state threshold to the signal precision is an artifact of the assumption of a diffuse prior for θ. I test the robustness of my comparative statics results to informative priors in Section 5 below.

of subversives. In turn, this means that as $\alpha \to \infty$ all regimes with $\theta \geq 0$ will survive. To see this, note for large α solutions to equation (26) are approximately the same as solutions to

$$\mathbb{1}\{\theta^*_\alpha - x^*_\alpha > 0\} = -\frac{p}{1-p} c(\theta^*_\alpha - x^*_\alpha) \tag{28}$$

The only solution to equation (28) is $\theta^*_\alpha - x^*_\alpha = 0$. So as $\alpha \to \infty$, solutions to equation (26) approach zero too. From equation (27) we now know that $\theta^*_\alpha \to 0^+$. Therefore, the signal-jamming technology is effective when the precision α is large enough. For large α the regime's threshold θ^*_α is less than the Morris-Shin level of $\theta^*_{\text{MS}} = 1-p$ and the regime's survival probability is correspondingly higher.

Intuition for the result. Why does a higher α help regimes? First, from (23) for any finite α and hence any $\gamma > 0$ there is an interval of regimes $[\theta^*, \theta^{**})$ who take actions that lead them to imitate the signal mean of a higher type of regime θ^{**} that is intrinsically harder to overthrow than they are. Second, with high α there is a large density of citizens near the signal mean (wherever that happens to be). Since an interval of regimes is able to imitate the signal mean of a higher type of regime, it is rational for any individual, when contemplating the beliefs of others, to assign relatively high probability to the event that they have signals near this artificially high mean $\theta^{**} > \theta^*$. At the margin this makes any individual less likely to attack and so the aggregate mass S who do is relatively low.[15] Two things are critical to this argument. First, different types of regimes must take different actions so that there is uncertainty about the amount by which individuals should discount signals: if all regimes took the same action, it would be easy to undo. And second, as discussed below, individuals must be imperfectly coordinated.

From a technological standpoint, signals may be intrinsically precise (of high quality). But this does not necessarily translate into reduced posterior uncertainty for individuals. The direct effect of higher α is to reduce posterior uncertainty, but there is also an indirect effect through the regime's hidden actions. Suppose the regime's policy was linear, $a(\theta) = a_0(\alpha) + a_1(\alpha)\theta$ for some coefficients $a_0(\alpha), a_1(\alpha)$ (this can't be true in equilibrium, but it's instructive nonetheless). If so, citizens would have normal posteriors with precision $\alpha[1+a_1(\alpha)]^2$. Then if $\alpha \to \infty$ but $a_1(\alpha) \to -1$ sufficiently fast, the signals x_i have no local information about θ even when α is large. The example with constant marginal cost has slope coefficient exactly -1 whenever a regime manipulates and so in this case signals are locally uninformative about θ.

3.3 Signal-jamming: the general case

The effectiveness of signal-jamming when α is large extends to the case of general convex costs:

[15] This intuition is not quite complete: there is a second effect. Note $\gamma \to 0$ as $\alpha \to \infty$ so $\theta^{**} \to \theta^*$ as $\alpha \to \infty$ and the interval of pooled types also shrinks as α becomes large. For the intuition in the text to be the dominant effect it also needs to be the case that $\gamma \to 0$ sufficiently slowly relative to the speed at which the density of signals at the signal mean increases. For large α, the rate at which the density increases is of order $\sqrt{\alpha}$ while for large α the constant γ is of order $\sqrt{\log(\alpha)/\alpha}$ which decreases slowly relative to the rate at which $\sqrt{\alpha}$ increases.

THEOREM 2. As the signal precision $\alpha \to \infty$ the limiting equilibrium thresholds and hidden action profile are

$$\lim_{\alpha\to\infty} \theta^*_\alpha = 0^+, \qquad \lim_{\alpha\to\infty} x^*_\alpha = 0^+, \quad \text{and} \quad \lim_{\alpha\to\infty} a_\alpha(\theta) = 0^+ \quad \text{for all } \theta$$

If in addition the cost of information manipulation is strictly convex, $C''(a) > 0$ all a, then as $\alpha \to 0^+$ the limiting equilibrium thresholds and hidden action profile are

$$\lim_{\alpha\to 0^+} \theta^*_\alpha = 1^-, \qquad \lim_{\alpha\to 0^+} x^*_\alpha = +\infty, \quad \text{and} \quad \lim_{\alpha\to 0^+} a_\alpha(\theta) = 0^+ \quad \text{for all } \theta$$

So for high enough α the signal-jamming technology is maximally effective. Moreover, there is a partial converse.[16] If costs are strictly convex then for low enough α the signal-jamming technology is ineffective in that $\theta^* > \theta^*_{\text{MS}} = 1-p$. If so, regimes would want to be able to credibly commit to not use the signal-jamming technology.

To characterize signal-jamming in the general case, rearrange (23) to get an alternative implicit characterization of the hidden actions

$$a(\theta) = x^* - \theta + \sqrt{\frac{2}{\alpha}\log\left(\frac{\sqrt{\alpha}\phi(0)}{C'[a(\theta)]}\right)}, \qquad \theta \geq \theta^* \tag{29}$$

which is (23) generalized to arbitrary convex costs but at the expense of losing the closed-form solution. By implicitly differentiating with respect to θ, it's possible to show that $a'(\theta) \geq -1$ with strict equality if $C''(a) > 0$ (so the signal mean $\theta + a(\theta)$ is increasing in θ, strictly if costs are strictly convex). But as $\alpha \to \infty$, regimes that manipulate have $a'(\theta) \to -1$ so that signals are locally uninformative. Moreover, since generally $a(\theta^*) > 0$, the signal mean for a regime that intervenes is strictly larger than θ^* so that signal-jamming is effective as α becomes sufficiently large because each individual worries about a large number of others drawing signals that suggest the regime is not going to be overthrown.

Numerical examples. With general cost functions the model cannot be solved analytically. Figure 3 shows θ^*_α as a function of precision α under the assumption that $C(a) := 0.5a^2$ for three levels of p. The higher the individual opportunity cost p, the lower the threshold and the thresholds are decreasing in the signal precision. In these examples, the speed of convergence to the limit is faster if p is high and slower if p is low. Regimes that inhabit a world where the individual cost of subversion p is high may benefit most from a given increase in α.

Interpretation. These results suggest that a regime's less overt propaganda apparatus (pressure exerted on editors, generals forced to stand on parade, etc) will be more useful when individuals are receiving signals that are of sufficiently high intrinsic quality (from a technological standpoint).

[16] As discussed in Section 3.1, if the marginal cost at zero is too large $C'(0) > \bar{c} := \sqrt{\alpha}\phi(0)$, then the cost of information manipulation is so high that the model reduces to the standard Morris-Shin game. When we take $\alpha \to \infty$ this bound does not matter. When we take $\alpha \to 0^+$ this bound will be violated. Consequently, the second part of Theorem 2 deals only with the case of strictly convex costs.

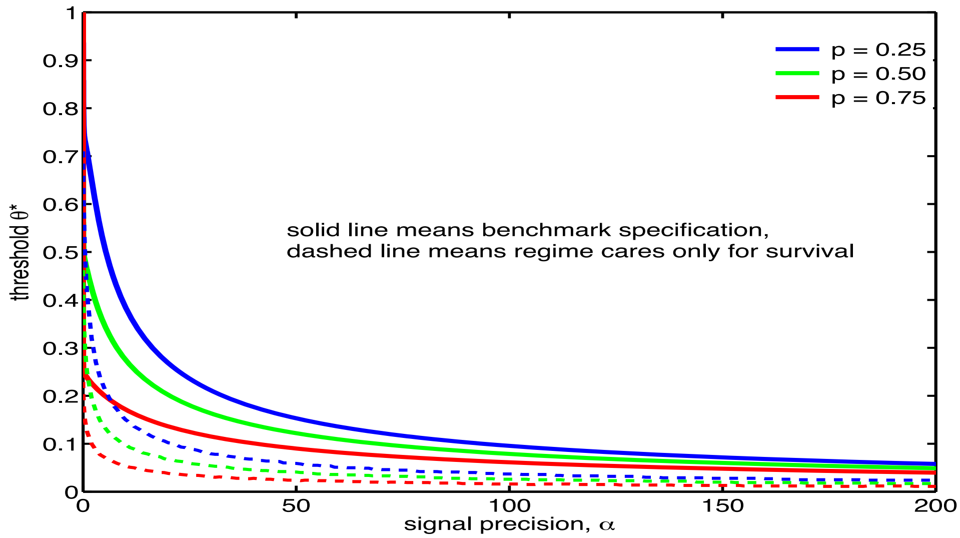

Figure 3: In the Morris-Shin game, thresholds are $\theta^*_{\mathsf{MS}} = 1 - p$ all α. With information manipulation, large α implies $\theta^*_\alpha \to 0^+$, so in the limit all 'fragile' regimes with $\theta \in [0, 1)$ survive. Solid lines show the benchmark specification while dashed lines show the alternative specification where the regime cares only about survival (see Section 4 below). In either case, better information increases a regime's likelihood of survival. All calculations use $C(a) := 0.5a^2$.

In equilibrium signals may be very uninformative, but that is precisely because the regime is co-opting the technology to its own ends.[17] A regime will want to exert a strong influence over the media when the signal precision is high enough.

Role of imperfect coordination. Disaggregated information is essential to the regime's ability to benefit from higher signal precision. To see this, suppose that citizens were perfectly coordinated and able to act as a large agent of mass one who could force regime change for all $\theta < 1$. Suppose also that this mass receives one signal x with precision α that is contaminated by the regime's action. Finally, suppose that costs are strictly convex. Then, as in the benchmark model, for high precision the regime receives almost no benefit from taking a costly action and $a(\theta) \to 0$ for all θ as $\alpha \to \infty$. But if citizens are coordinated, they then know that $x \to \theta$ and so attack if and only if $x = \theta < 1$. If citizens are perfectly coordinated, all fragile regimes with $\theta \in [0, 1)$ are wiped out. But if citizens are imperfectly coordinated all fragile regimes with $\theta \in [0, 1)$ survive. See Appendix B for more details.

3.4 Equilibrium beliefs

Regimes benefit when α is sufficiently high. Is this because in equilibrium citizens somehow fail to filter out the information manipulation? Yes and no: 'no' in that for high enough α all citizens

[17] A clear example of a regime co-opting new technologies for propaganda purposes is the sponsored diffusion of the cheap *Volksempfänger* radio set in 1930s Germany. By 1939, 70% of households owned a set — the highest proportion in the world at the time (Zeman, 1973, 34-62).

discount their signals, but 'yes' in that they typically do not discount *enough*.

To see this, first suppose that a citizen with signal x_i was naive and believed she lived in the Morris-Shin world where no manipulation takes place. Her posterior expectation for θ would equal x_i. Let $D(x_i)$ denote the difference between the rational expectation of θ and the naive expectation

$$D(x_i) := \int_{-\infty}^{\infty} \theta\pi(\theta|x_i)d\theta - x_i$$

If $D(x_i) > 0$, the regime's propaganda has successfully induced this citizen to believe that θ is relatively high. Otherwise, if $D(x_i) < 0$, this citizen believes θ is relatively low. The average $\int D(x_i)f(x_i|\theta, a(\theta))dx_i = -a(\theta) \leq 0$, so on average the population generally discounts signals.

Moreover, if α is sufficiently high the distribution of x_i is almost degenerate for any θ so the average D can only be non-positive if $D(x_i) \leq 0$ for all x_i. Only if α is sufficiently low, such that the conditional expectation of θ is almost independent of x_i, can there exist some x_i who are 'fooled' in equilibrium, $D(x_i) > 0$. The top panel of Figure 4 illustrates. So for high enough α we can agree that all citizens discount their signals.

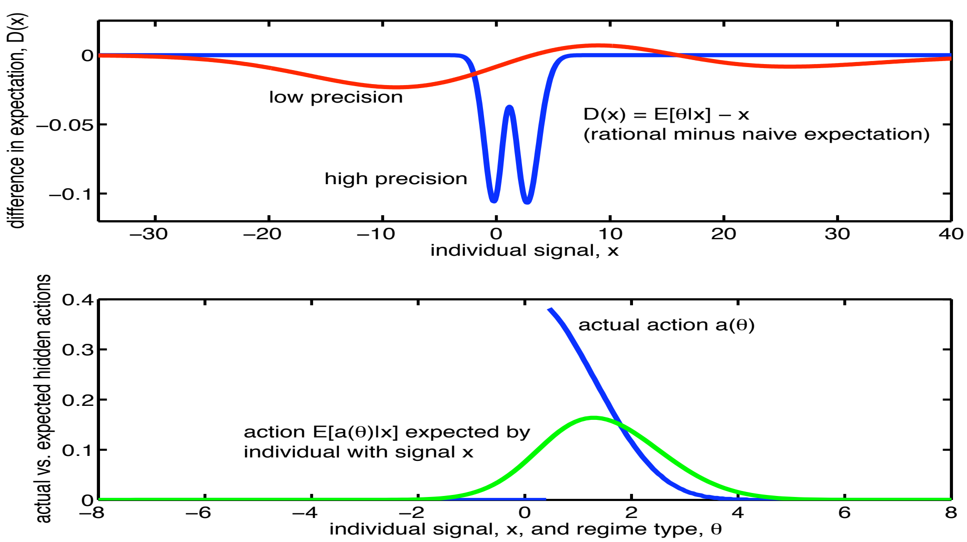

Figure 4: If precision is high enough, all citizens discount their signals: the difference between the rational and naive expectations is $D(x_i) = \int \theta\pi(\theta|x_i)d\theta - x_i < 0$. But if precision is low enough, some citizens believe the regime's propaganda and have $D(x_i) > 0$ (top panel). While citizens on average discount, they may not discount their signals enough. There are θ such that $a(\theta) \geq \int a(\theta)\pi(\theta|x_i)d\theta$ for every x_i (bottom panel).

But this does not mean they discount their signals *enough*. The bottom panel of Figure 4 shows the relationship between $a(\theta)$, which is the amount by which citizens should discount their information, and $\int a(\theta)\pi(\theta|x_i)d\theta$, which is the amount that someone with x_i does discount their signal. Critically, there are regimes with $\theta \geq \theta^*$ that engage in more manipulation in equilibrium than is believed by *any* citizen. Citizens 'under-discount' the manipulation done by fragile regimes just above the threshold θ^*. Similarly, since $a(\theta) = 0$ for $\theta < \theta^*$ and $a(\theta) \to 0$ for large θ, most citizens 'over-discount' the manipulation done by either very weak or very strong regimes.

The model that delivers these results is stylized and it's natural to question its robustness. To investigate this, Section 4 extends the model to consider regimes that only care about survival. Regimes still benefit from better information. Alternative information structures are covered in Section 5: for some but not all of these structures, regimes benefit from better information.

4 Regime cares only about survival

Let the payoff to a regime be $B(S, \theta) - C(a)$ where

$$B(S, \theta) = \mathbb{1}\{\theta \geq S\} v(\theta) \tag{30}$$

so that a regime cares about S only to the extent that a high S makes it harder to survive. I consider two specifications for $v(\theta)$. In the first, $v(\theta) = \theta$ so that as in the benchmark model the (gross) value a regime attaches to survival is perfectly correlated with its 'strength' as measured by its ability to withstand a mass S of given size. In the second specification, $v(\theta) = \overline{v} > 0$ so that the value a regime attaches to survival is independent of its strength (and is common knowledge).

Characterization of equilibrium. With these payoffs, only fragile regimes with $\theta \in [0, 1)$ could possibly have an incentive to manipulate. The mass of subversives is $S(\theta, a) = \Phi[\sqrt{\alpha}(x^* - \theta - a)]$ and if an action $a(\theta) > 0$ is chosen, it is set at just the right level to ensure that $\theta = S(\theta, a)$. Specifically

$$a(\theta) = x^* - \theta - \frac{1}{\sqrt{\alpha}} \Phi^{-1}(\theta), \qquad \theta \in [\theta^*, \theta^{**}) \tag{31}$$

and zero otherwise. Any higher an action would incur a cost for no benefit. Given a signal threshold x^*, the upper boundary $\theta^{**} \in [\theta^*, 1)$ is the unique solution to $\theta^{**} = \Phi[\sqrt{\alpha}(x^* - \theta^{**})]$. Given x^*, equation (31) determines the action policy and the threshold below which the regime is overthrown is determined by the indifference condition $v(\theta^*) = C[a(\theta^*)]$. The right hand side of (31) implies that $C[a(\theta)]$ is continuous and strictly decreasing in θ. Now use the intermediate value theorem and that $v(\theta)$ is continuous and weakly increasing in θ for either specification ($v(\theta) = \theta$ or $v(\theta) = \overline{v}$). This implies a unique solution $\theta^* \in [0, \theta^{**})$. The hidden action policy (31) can then be used to calculate posterior beliefs for a citizen with signal x_i as in equation (16) and this implicitly determines a threshold x^* as a function of θ^*. An equilibrium is calculated by simultaneously solving these two conditions for x^*, θ^*.

4.1 Regime type and value of survival are the same.

Let $v(\theta) = \theta$ and specialize to $C(a) = ca$ when α is large. Let x^*, θ^* and $a(\theta)$ denote equilibrium objects for the benchmark economy and let $\hat{x}^*, \hat{\theta}^*$ and $\hat{a}(\theta)$ denote the same objects in the economy where the regime cares only about survival.

From (23) we have approximately $a(\theta) = \max[0, x^* - \theta]$ when α is large. But from (31) we also have approximately $\hat{a}(\theta) = \max[0, \hat{x}^* - \theta]$ when α is large. So if $x^* = \hat{x}^*$, then $a(\theta) = \hat{a}(\theta)$ for each θ. The thresholds $\theta^*, \hat{\theta}^*$ are determined similarly: θ^* is determined by $\theta^* = ca(\theta^*)$ from (27) when $\alpha \to \infty$ and if the regime cares only about survival $\hat{\theta}^* = c\hat{a}(\hat{\theta}^*)$ for any α. So again if $x^* = \hat{x}^*$, then $a(\theta^*) = \hat{a}(\hat{\theta}^*)$ so that $\theta^* = \hat{\theta}^*$. But for either payoff specification for the regime the signal threshold is uniquely determined by the citizen indifference condition (18) once the hidden actions and state threshold are known. So for large α indeed $x^* = \hat{x}^*$ and the entire equilibrium will be approximately the same under either specification.

To further check robustness, I have solved the model numerically with strictly convex costs and when α is small. Figure 3 shows $\theta^*, \hat{\theta}^*$ as functions of α for various levels of p. The qualitative similarity between the model with the benchmark payoffs and the alternative payoffs is clear even for relatively low α.

4.2 Regime type and value of survival are independent.

A similar result obtains for the other specification where $v(\theta) = \overline{v}$ all θ. Hidden actions are still given by (31) and so for large α are approximately $a(\theta) = \max[0, x^* - \theta]$. The threshold θ^* is now determined by $\overline{v} = C[a(\theta^*)]$. If we again specialize to $C(a) = ca$ an analogous argument shows that as $\alpha \to \infty$ we have $x^* \to C^{-1}(\overline{v}) = \overline{v}/c > 0$ and $\theta^* \to 0$. As with the benchmark payoffs the regime benefits for large enough α.

5 Alternative information structures

Until now, improvements in the quality of information were modeled as increases in the precision α. But in the benchmark model there is only one kind of information and it is affected by the regime's action. Section 5.1 allows citizens to have clean sources of information *uncontaminated* by the regime's actions. Section 5.2 changes the technology available to the regime, giving it the ability to manipulate the precision of signal information rather than the signal mean.

5.1 Uncontaminated information

I first consider the case where citizens are endowed with a second private signal independent of the regime's a and then turn to the case of a public signal. This second exercise connects the model to recent papers in the global games literature regarding the differential effects of private and public information on equilibrium uniqueness.

Heterogeneous informative priors. Let citizens have two idiosyncratically noisy signals

$$\begin{aligned} x_i &= \theta + a + \varepsilon_{x,i} \\ z_i &= \theta \qquad + \varepsilon_{z,i} \end{aligned} \tag{32}$$

where $\varepsilon_{z,i}$ is normal noise with precision α_z independent of $\varepsilon_{x,i}$ which has precision α_x. The z_i signals can be treated as heterogeneous priors and unlike x_i are not contaminated by the regime's choice of a. This is equivalent to idiosyncratically noisy information on the hidden action, say $y_i = a + \xi_i$ where $y_i := x_i - z_i$ and $\xi_i := \varepsilon_{x,i} - \varepsilon_{z,i}$.

A monotone equilibrium consists of hidden actions $a(\theta)$, state threshold θ^* and signal thresholds $x^*(z_i)$ such that a citizen with x_i, z_i subverts if $x_i < x^*(z_i)$. Unfortunately, it is difficult to provide an analytic characterization of the equilibrium with both contaminated and uncontaminated private information, so the discussion below merely provides examples. The key finding is that an exogenous increase in the precision of *uncontaminated* information α_z may still reduce θ^*.

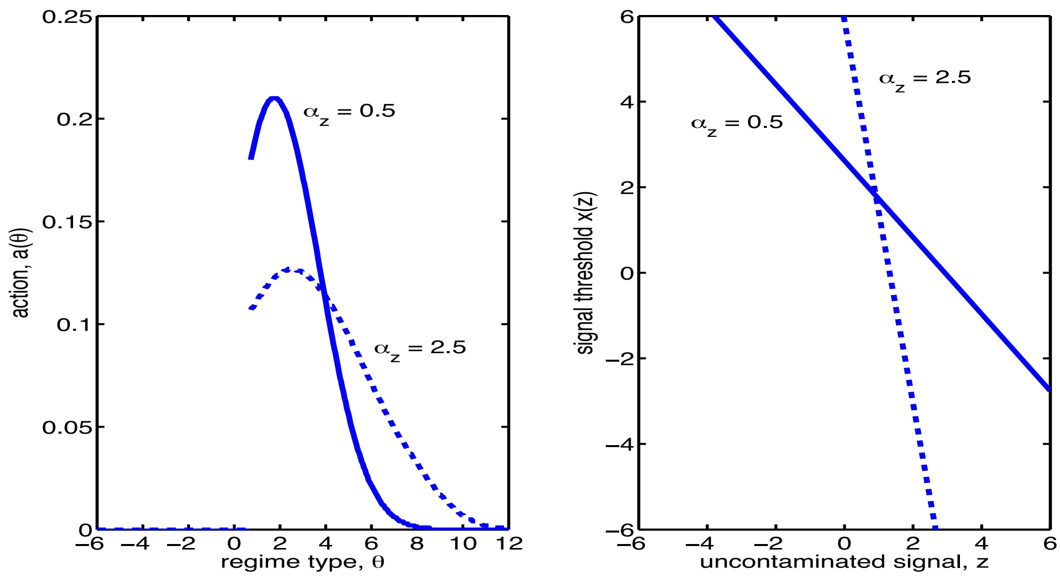

Figure 5: As precision α_z of uncontaminated private information increases, regimes near θ^* take smaller actions $a(\theta)$ but diminishing returns do not set in at such low values of θ. Citizens give more weight to their uncontaminated information so $x^*(z_i)$ is steeper: for low z_i it takes a higher contaminated x_i to induce subversion.

Figure 5 shows the actions $a(\theta)$ and the thresholds $x^*(z_i)$ for $\alpha_z = \alpha_x = 0.5$ and for $\alpha_z = 2.5$. As the uncontaminated information increases in quality, regimes with types larger than θ^* take smaller actions. But when α_z is high, diminishing returns do not set in so quickly: $a(\theta)$ is noticeably positive for more θ. The right hand panel shows the corresponding thresholds $x^*(z_i)$; these are decreasing because if an individual gets a low z_i it takes a high x_i to induce subversion. As α_z increases, $x^*(z_i)$ becomes steeper so that the z_i are weighed more heavily and it takes, for example, an even bigger x_i to compensate for a low z_i.

Figure 6 shows θ^* as a function of the uncontaminated precision α_z for various opportunity costs p. The θ^* are lower than the Morris-Shin levels $1 - p$ so that signal-jamming is effective and are decreasing in the precision α_z of the uncontaminated private information. But compared to the model with only contaminated information, the effects of an increase in α_z on θ^* are small.

These examples are only suggestive of what can happen in equilibrium. But it is clear that introducing uncontaminated private information does not necessarily overturn the results of Section 3. A *clean* information revolution may still increase the survival probability of a regime.

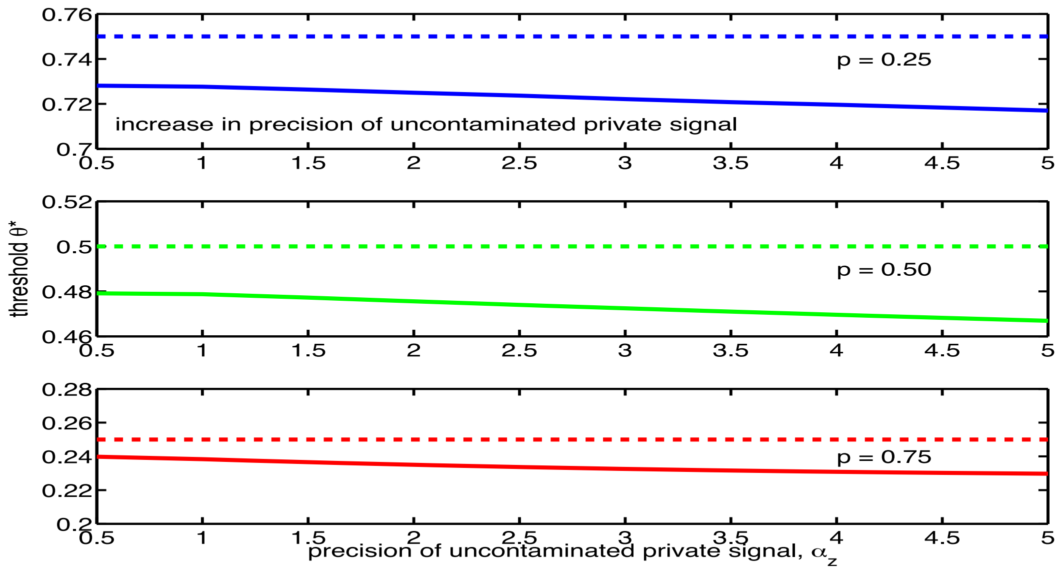

Figure 6: Thresholds θ^* as function of precision of uncontaminated private information α_z. In these examples a regime may benefit even from an increase in the quality of 'clean' information. Equilibrium thresholds are relatively insensitive to α_z.

Common informative priors/public signal. Let citizens have a public signal

$$\begin{array}{rcllll} x_i & = & \theta + a & + & \varepsilon_{x,i} \\ z & = & \theta & + & \varepsilon_z \end{array} \tag{33}$$

The realization of z is common knowledge and again is an uncontaminated signal of θ.

If no hidden actions are possible, the model reduces to a setup studied by Angeletos and Werning (2006), Hellwig (2002), Metz (2002), and Morris and Shin (2000, 2003) and others. For each z, a monotone equilibrium is a threshold $x^*(z)$ such that individuals subvert if $x_i < x^*(z)$ and a $\theta^*(z)$ such that regimes are overthrown if $\theta < \theta^*(z)$. It is well known that if public information is too precise relative to private information, there may be multiple monotone equilibria [see, e.g., Hellwig (2002) and Morris and Shin (2003, 2004)]. If public information is too precise, there is 'approximate' common knowledge of θ. Hellwig (2002) derived a sufficient condition for a unique monotone equilibrium in a game of this kind, namely $\alpha_z/\sqrt{\alpha_x} < \sqrt{2\pi}$. In the discussion that follows, I assume that this is satisfied.

What happens if the precision of private information increases? As $\alpha_x \to \infty$ for given α_z (or as $\alpha_z \to 0$ for given α_x) the public signal becomes uninformative and we revert to the Morris-Shin game with $x^*(z) \to \theta^*(z)$ and $\theta^*(z) \to 1-p$ independent of the realization of the public signal z. If the quality of private information is sufficiently good, the public information z is irrelevant.

Metz (2002) characterized the direction from which $\theta^*(z)$ converges to $1-p$ as $\alpha_x \to \infty$. If the parameters p or z are favorable to the regime, then $\theta^*(z) \nearrow 1-p$ but if the parameters p or z are unfavorable to the regime, $\theta^*(z) \searrow 1-p$. Why? Both x_i and z are informative about θ, but only x_i is informative about the role of coordination. Also, it is common knowledge that signals are bunched around θ and that every citizen gives weight to their idiosyncratic signal in proportion to its quality. Now consider an economy with high z (which suggests the regime is going to be difficult to beat, since high z is correlated with high θ). For moderate α_x, citizens will give some weight to this public signal and will be less inclined to engage in subversion. So for moderate precision, $\theta^*(z)$ is low and the ex ante survival probability of the regime is high. But as α_x increases, the influence of the high realized z diminishes because everybody knows that everybody gives less weight to z when α_x increases. In the limit, only the opportunity cost p matters and $\theta^*(z) \nearrow 1-p$. I refer to this as the *coordination effect* from increasing idiosyncratic signal precision.

With hidden actions the coordination effect is dominant for low α_x and $\theta^*(z)$ is increasing in α_x if p or z is high but decreasing in α_x if p or z is low. But the coordination effect is limited: when α_x is large, people ignore z and the effects of further increase in α_x are almost nil. Now recall the basic hidden action model with no public signal: the state threshold approached zero as the idiosyncratic signal precision became large. So we expect for high α_x, the existence of a public signal is almost immaterial and the *hidden action effect* of information manipulation is dominant.

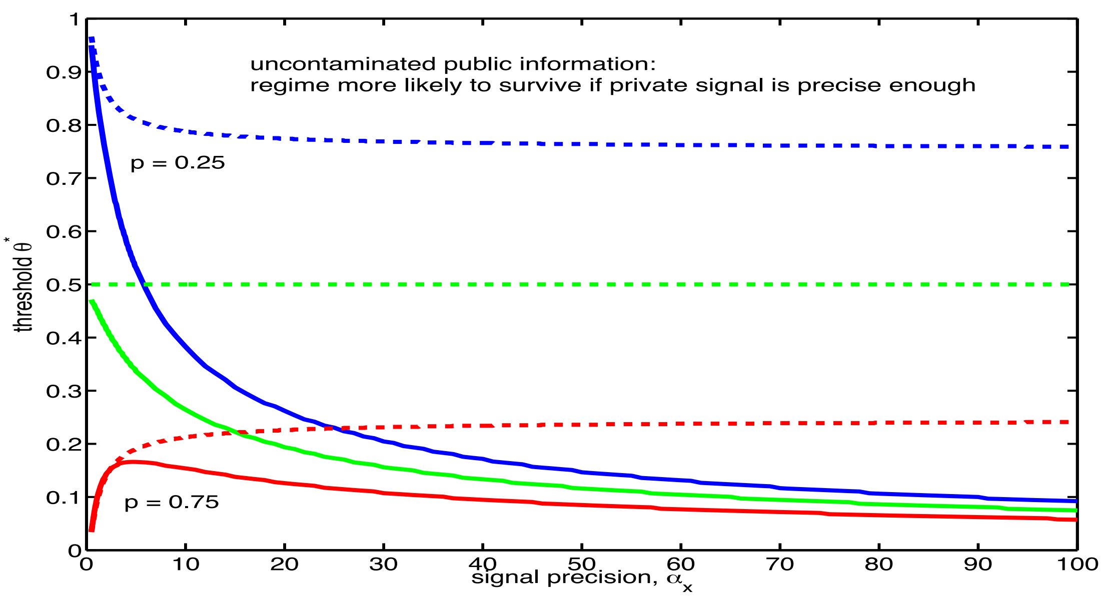

Figure 7: Thresholds θ^* as functions of α_x. Dashed lines show the thresholds when there are no hidden actions. Solid lines show θ^* when hidden actions contaminate the private information. If primitives are favorable (say $p = 0.75$), the coordination effect identified by Metz (2002) and the hidden action effect pull in opposite directions, giving rise to a 'hump-shaped' function. If primitives are unfavorable to the regime (say $p = 0.25$), both forces drive θ^* down.

Figure 7 suggests that this intuition is correct. When primitives are favorable to the regime, the coordination effect and the hidden action effect pull in opposing directions. The coordination effect tends to drive the equilibrium threshold up and to reduce the survival probability of the regime. But for high α_x the coordination effect is irrelevant while the hidden action effect still matters. In this case, the thresholds have a 'hump-shaped' look and asymptote to zero. But if primitives are unfavorable to the regime, both the coordination and hidden action effects are in the regime's favor and reinforce each other. In this case, the threshold is monotone decreasing and asymptotes to zero. If primitives are unfavorable to the regime, there is a large benefit from information manipulation (this contrasts with the benchmark model where regimes with favorable primitives benefitted more from a given increase in α_x).

Uncontaminated public information does not undo the central message of the hidden actions model. For high α_x signal-jamming is effective, $\theta^* < \theta^*_{\text{MS}} = 1 - p$. Better private information may increase the regime's ex ante survival probability. But perhaps this is not the most interesting comparative static. What if the precision of the *public* signal increases? Then we run into the problem of multiplicity. For given α_x, the inequality $\alpha_z/\sqrt{\alpha_x} < \sqrt{2\pi}$ will eventually be violated and there are multiple monotone equilibria. In this case, we lose the ability to draw sharp conclusions. What if both the precisions of the public and private information increase together? Then if the bound on the relative precision of the public information is satisfied the analysis goes through essentially as above. At a given level of the relative precision $\alpha_z/\sqrt{\alpha_x}$, the threshold may be increasing or decreasing (depending on the relative strengths of the coordination effect and the hidden action effect), but as the ratio $\alpha_z/\sqrt{\alpha_x}$ becomes small the hidden action effect dominates.

5.2 Manipulating signal precision

Until now, signal manipulation entered in an additive way: $x_i := \theta + a + \varepsilon_i$. With this specification the action shifts the *mean* of the signal distribution without (directly) influencing the precision. Now let signals be $x_i := \theta + \varepsilon_i$ with the precision of ε_i given by

$$\beta(a) := \alpha\left(\frac{1}{2} + \Phi(a)\right), \qquad \alpha > 0 \tag{34}$$

With this specification, negative actions $a < 0$ are possible and we need to reinterpret the cost function C. When $a > 0$, the regime sets a precision $\beta(a) > \alpha$ and when $a < 0$, the regime sets a precision $\beta(a) < \alpha$. Accordingly, I refer to α as the *intrinsic precision* of private information.

In equilibrium, hidden actions $a(\theta)$ are characterized by the first order necessary condition

$$C'(a) = (\theta - x^*)\sqrt{\beta(a)}\phi\left[\sqrt{\beta(a)}(\theta - x^*)\right]\frac{\alpha}{2}, \qquad \theta \geq \theta^* \tag{35}$$

Actions are zero for $\theta < \theta^*$ before jumping discontinuously at θ^*. If $x^* > \theta^*$, the jump is down but if $x^* < \theta^*$ the jump is up. Intuitively, if the regime has intermediate type $\theta \in [\theta^*, x^*)$ then it worsens the signal noise relative to α (it *muddies* the signal). Otherwise, if $\theta > x^*$ then the regime finds it worthwhile to *clarify* its position of strength and so increase the signal precision relative to α. Figure 8 illustrates with parameters chosen so $x^* > \theta^*$ and the profile jumps down at θ^*.

Transparency. In related work Bannier and Heinemann (2005) and Heinemann and Illing (2002) study higher signal precision (more 'transparency') in the Morris-Shin game by allowing the regime to choose the individuals' signal precision to minimize the probability of a successful attack. But in their formulation the regime's choice of precision does not depend on its type θ and so when forming beliefs individuals can take this precision as exogenous. In Figure 8, a regime reduces transparency in this sense if $\theta \in [\theta^*, x^*)$ and increases transparency if $\theta > x^*$.

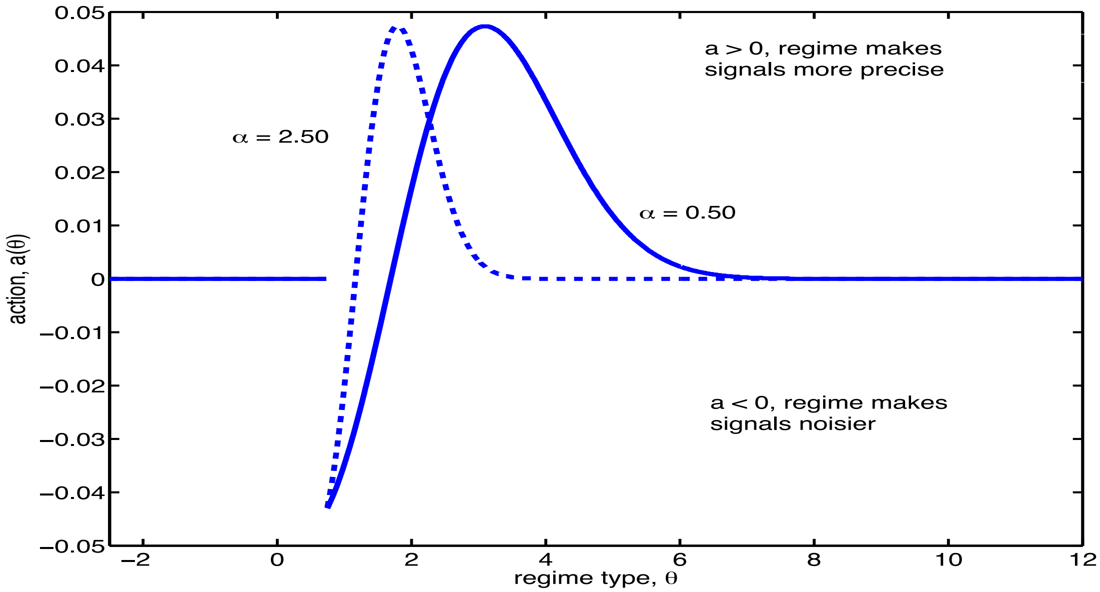

Figure 8: Equilibrium hidden actions $a(\theta)$ when regime can manipulate signal precision. For intermediate θ it may be optimal for the regime to 'muddy' the signal, $a(\theta) < 0$, relative to the benchmark α. For high θ it is optimal for the regime to clarify its strength, $a(\theta) > 0$.

6 Final remarks

The coordination games studied in this paper are deliberately stylized so as to focus attention on the use of information manipulation as a tool for influencing equilibrium outcomes. In keeping things simple, I have abstracted from issues that might play a role in the development of a more complete model of information, coordination and regime change. For example, information accumulation over time, non-simultaneous moves by groups of citizens and strategic communication between individuals might all be important components of a richer analysis. Moreover, for some applications it might be desirable to have several 'large' players with competing interests over equilibrium outcomes (for example, if one large player has an interest in the status quo while a rival large player seeks its overthrow). In such situations, individuals beliefs and equilibrium outcomes would depend on the interactions between several hidden actions each of which is trying to push beliefs in a different direction.

A Proofs

Proof of Proposition 1. (Morris-Shin): Let $\hat{x}, \hat{\theta}$ denote candidate thresholds. Posterior beliefs of a citizen with x_i facing $\hat{\theta}$ are $\Pr(\theta < \hat{\theta}|x_i) = \Phi[\sqrt{\alpha}(\hat{\theta} - x_i)]$. A citizen with x_i will subvert if and only if $\Phi[\sqrt{\alpha}(\hat{\theta} - x_i)] \geq p$. This probability is continuous and strictly decreasing in x_i, so for each $\hat{\theta}$ there is a unique signal for which a citizen is indifferent. Similarly, if the regime faces threshold $\hat{x}$ the mass of subversives is $\hat{S}(\theta) = \Phi[\sqrt{\alpha}(\hat{x} - \theta)]$. A regime θ will not be overthrown if and only if $\theta \geq \Phi[\sqrt{\alpha}(\hat{x} - \theta)]$. The probability on the right hand side is continuous and strictly decreasing in θ, so for each $\hat{x}$ there is a unique state for which a regime is indifferent. The Morris-Shin thresholds $x^*_{\text{MS}}, \theta^*_{\text{MS}}$ simultaneously solve these best response conditions as equalities. It is straightforward to verify that there is only one solution to these equations and that $\theta^*_{\text{MS}} = 1 - p$ all α. □

Proof of Lemma 1. Fix $\hat{x} \in \mathbb{R}$. (i) From the envelope theorem $W'_{\hat{x}}(\theta) = 1 - S'_{\hat{x}}(\theta + a) > 1$ all θ since $S'_{\hat{x}}(\theta + a) < 0$. Since $S_{\hat{x}} \geq 0$ and $C \geq 0$ we know $W_{\hat{x}}(\theta) < 0$ all $\theta < 0$. Similarly, $W_{\hat{x}}(1) > 0$. So by the intermediate value theorem there is a unique $\theta_{\hat{x}} \in [0, 1)$ such that $W_{\hat{x}}(\theta_{\hat{x}}) = 0$. And since $W'_{\hat{x}}(\theta) > 1$ the regime is overthrown if and only if $\theta < \theta_{\hat{x}}$. (ii) Since positive actions are costly, the regime takes no action for $\theta < \theta_{\hat{x}}$. Otherwise, for $\theta \geq \theta_{\hat{x}}$, the actions of the regime solve $a_{\hat{x}}(\theta) \in \text{argmax}_{a \geq 0} [\theta - S_{\hat{x}}(\theta + a) - C(a)]$. The first order necessary condition for interior solutions can be written $C'(a) = \sqrt{\alpha}\phi[\sqrt{\alpha}(\hat{x} - \theta - a)]$. Using $\phi(w) := \exp(-w^2/2)/\sqrt{2\pi}$ all $w \in \mathbb{R}$, the first order condition may have zero, one or two solutions for each θ. If for given θ there are zero (interior) solutions, then $a_{\hat{x}}(\theta) = 0$. If for given θ there are two solutions, one of them can be ruled out by the second order sufficient condition $-\alpha\phi'[\sqrt{\alpha}(\hat{x} - \theta - a)] - C''(a) < 0$. Using the property $\phi'(w) = -w\phi(w)$ all $w \in \mathbb{R}$ shows that if there are two solutions to the first order condition, only the 'higher' of them satisfies the second order condition. Therefore for each θ there is a single $a_{\hat{x}}(\theta)$ that solves the regime's problem. □

Proof of Lemma 2. For notational simplicity, write θ for the state threshold, x for an individual's signal and $P(\theta, x)$ for the probability an individual with x assigns to the regime's type being less that θ when the actions are $\hat{a} : \mathbb{R} \to \mathbb{R}_+$. From (15), this probability can be written

$$P(\theta, x) = \frac{A(\theta - x)}{A(\theta - x) + B(\theta, x)} \tag{36}$$

where

$$A(\theta - x) := \Phi[\sqrt{\alpha}(\theta - x)], \quad \text{and} \quad B(\theta, x) := \int_\theta^\infty \sqrt{\alpha}\phi[\sqrt{\alpha}(x - \xi - \hat{a}(\xi))]d\xi \tag{37}$$

Hence $P : \mathbb{R} \times \mathbb{R} \to [0, 1]$ is continuous in θ, x. (i) Since $A' > 0$ and $B_\theta < 0$, $P_\theta > 0$ for all θ, x. Moreover, since $A(-\infty) = 0$ and $B > 0$ we have $P(-\infty, x) = 0$ for all x. Similarly, since $\hat{a}(\xi) = 0$ for all $\xi < \theta$ as $\theta \to \infty$ we have $B(\theta, x) \to 1 - A(\theta - x)$ and since $A(\infty) = 1$ this means $B(\infty, x) = 0$ for all x. Therefore $P(\infty, x) = 1$ for all x. (ii) Differentiating (36) shows $P_x < 0$ if and only if $A'/A > -B_x/B$. Calculating the derivatives shows that this is equivalent to

$$H[\sqrt{\alpha}(x - \theta)] > -\frac{\int_\theta^\infty \phi'[\sqrt{\alpha}(x - \hat{y}(\xi))]d\xi}{\int_\theta^\infty \phi[\sqrt{\alpha}(x - \hat{y}(\xi))]d\xi} = \frac{\int_\theta^\infty \sqrt{\alpha}(x - \hat{y}(\xi))\phi[\sqrt{\alpha}(x - \hat{y}(\xi))]d\xi}{\int_\theta^\infty \phi[\sqrt{\alpha}(x - \hat{y}(\xi))]d\xi} \tag{38}$$

where $H(w) := \phi(w)/[1 - \Phi(w)] > 0$ denotes the standard normal hazard function for $w \in \mathbb{R}$, where $\hat{y}(\xi) := \xi + \hat{a}(\xi)$ is the mean of the signal distribution if $\xi \geq \theta$, and where the equality follows from $\phi'(w) = -w\phi(w)$ all w. Now define a density $\varphi(\xi|x) > 0$ by

$$\varphi(\xi|x) := \frac{\phi[\sqrt{\alpha}(x - \hat{y}(\xi))]}{\int_\theta^\infty \phi[\sqrt{\alpha}(x - \hat{y}(\xi'))]d\xi'}, \qquad \xi \in [\theta, \infty) \tag{39}$$

Then after a slight rearrangement of terms in (38), $P_x < 0$ if and only if

$$H[\sqrt{\alpha}(x - \theta)] - \sqrt{\alpha}(x - \theta) > \sqrt{\alpha}\left[\theta - \int_\theta^\infty \hat{y}(\xi)\varphi(\xi|x)d\xi\right] \tag{40}$$

Since the hazard function satisfies $H(w) > w$ all $w \in \mathbb{R}$ and $\alpha > 0$ it is sufficient that

$$\int_\theta^\infty \hat{y}(\xi)\varphi(\xi|x)d\xi \geq \theta \tag{41}$$

But since $\hat{y}(\xi) := \xi + \hat{a}(\xi)$, $\xi \geq \theta$, and $\hat{a}(\xi) \geq 0$, condition (41) is always satisfied. Therefore $P_x < 0$. The limit properties in x are established in parallel fashion to the limit properties in θ shown in part (i). $\square$

Proof of Lemma 3. Fix $\hat{x} \in \mathbb{R}$ and let $a(\theta - \hat{x})$ denote the associated hidden actions. Write $K(\hat{x}, x) := P(\theta_{\hat{x}}, x, \hat{x})$. (i) Immediate from Lemma 2. (ii) Analogous to (36) let $P(\theta_{\hat{x}}, x, \hat{x}) = A(\theta_{\hat{x}} - x)/[A(\theta_{\hat{x}} - x) + B(\theta_{\hat{x}}, x, \hat{x})]$ where $A : \mathbb{R} \to [0, 1]$ is defined as in (37) and where

$$B(\theta_{\hat{x}}, x, \hat{x}) := \int_{\theta_{\hat{x}}}^{\infty} \sqrt{\alpha}\phi[\sqrt{\alpha}(x - \theta - a(\theta - \hat{x}))]d\theta \tag{42}$$

Using this representation, we have that for $K(\hat{x}, x) = P(\theta_{\hat{x}} - \hat{x}, x - \hat{x}, 0)$ it is sufficient that $B(\theta_{\hat{x}}, x, \hat{x}) = B(\theta_{\hat{x}} - \hat{x}, x - \hat{x}, 0)$. From (42) and using the change of variables $\xi := \theta - \hat{x}$ we have

$$B(\theta_{\hat{x}}, x, \hat{x}) = \int_{\theta_{\hat{x}} - \hat{x}}^{\infty} \sqrt{\alpha}\phi[\sqrt{\alpha}(x - \hat{x} - \xi - a(\xi))]d\xi = B(\theta_{\hat{x}} - \hat{x}, x - \hat{x}, 0) \tag{43}$$

Therefore $K(\hat{x}, x) = P(\theta_{\hat{x}} - \hat{x}, x - \hat{x}, 0)$ too. $\square$

Proof of Lemma 4. The most *pessimistic* scenario for any citizen is that regimes are overthrown only if $\theta < 0$ and that regimes take the largest hidden actions that could be rational $\underline{a}(\theta) := C^{-1}(\theta)$ for $\theta \geq 0$ and zero otherwise. Let $\underline{P}(x) := \Pr[\theta < 0 \mid x,\ \underline{a}(\cdot)]$ denote the probability the regime is overthrown in this most pessimistic scenario. Lemma 2 holds for hidden actions of the form $\underline{a}(\theta)$ and implies $\underline{P}'(x) < 0$ all x and since $\underline{P}(-\infty) = 1$ and $\underline{P}(\infty) = 0$ by the intermediate value theorem there is a unique value, $\underline{x}$, finite, such that $\underline{P}(\underline{x}) = p$. For $x < \underline{x}$ it is (iteratively) strictly dominant for $s(x) = 1$. Similarly, the most *optimistic* scenario for any citizen is that regimes are overthrown if $\theta < 1$ and that regimes take the smallest hidden actions that could be rational $\overline{a}(\theta) := 0$. Let $\overline{P}(x) := \Pr[\theta < 1 \mid x,\ \overline{a}(\cdot)]$ denote the probability the regime is overthrown in this most optimistic scenario. A parallel argument establishes the existence of a unique value, $\overline{x}$, finite, such that $\overline{P}(\overline{x}) = p$. For $x > \overline{x}$ it is (iteratively) strictly dominant for $s(x) = 0$. $\square$

Proof of Lemma 5. All citizens with signals $x < \underline{x}$ have $s(x) = 1$ so the mass of subversives is at least $\Phi[\sqrt{\alpha}(\underline{x} - (\theta + a))]$. To acknowledge this, write the total mass of subversives as

$$S(\theta + a) = \Phi[\sqrt{\alpha}(\underline{x} - (\theta + a))] + \Delta(\theta + a) \tag{44}$$

for some function $\Delta : \mathbb{R} \to [0, 1]$. First consider the case $\Delta = 0$ where *only* citizens with $x < \underline{x}$ subvert the regime. From Lemma 1 there is a unique threshold $\theta_{\underline{x}} \in [0, 1)$ sustained by hidden action $a_{\underline{x}}(\theta_{\underline{x}}) \geq 0$ solving (13)-(14) such that the regime is overthrown if $\theta < \theta_{\underline{x}}$. Now consider the case $\Delta > 0$ where *some* citizens with signals $x \geq \underline{x}$ also subvert the regime. The proof that the regime is overthrown for at least all $\theta < \theta_{\underline{x}}$ is by contradiction. Suppose that when $\Delta > 0$ regime change occurs for all $\theta < \tilde{\theta}$ for some $\tilde{\theta} \leq \theta_{\underline{x}}$. A marginal regime $\tilde{\theta}$ must be indifferent between being overthrown and taking the outside option, so this threshold satisfies $\tilde{\theta} = S(\tilde{\theta} + \tilde{a}) + C(\tilde{a})$ where $\tilde{a} \geq 0$ is the optimal action for the marginal regime $\tilde{\theta}$. Then observe

$$\begin{aligned}
\theta_{\underline{x}} &= \Phi[\sqrt{\alpha}(\underline{x} - (\theta_{\underline{x}} + a_{\underline{x}}(\theta_{\underline{x}})))] + C(a_{\underline{x}}(\theta_{\underline{x}})) \\
&\leq \Phi[\sqrt{\alpha}(\underline{x} - (\theta_{\underline{x}} + a))] + C(a), \\
&< \Phi[\sqrt{\alpha}(\underline{x} - (\theta_{\underline{x}} + a))] + \Delta(\tilde{\theta} + \tilde{a}) + C(a), \qquad \text{for any } a \geq 0
\end{aligned}$$

where the first inequality follows because $a_{\underline{x}}(\theta)$ minimizes $\Phi[\sqrt{\alpha}(\underline{x}-(\theta+a))]+C(a)$ and where the second inequality follows from $\Delta>0$. Taking $a=\tilde{a}\geq 0$ we then have

$$\begin{aligned}\theta_{\underline{x}} &< \Phi[\sqrt{\alpha}(\underline{x}-(\theta_{\underline{x}}+\tilde{a}))]+\Delta(\tilde{\theta}+\tilde{a})+C(\tilde{a})\\ &= \Phi[\sqrt{\alpha}(\underline{x}-(\theta_{\underline{x}}+\tilde{a}))]+\Delta(\tilde{\theta}+\tilde{a})+C(\tilde{a})+\Phi[\sqrt{\alpha}(\underline{x}-(\tilde{\theta}+\tilde{a}))]-\Phi[\sqrt{\alpha}(\underline{x}-(\tilde{\theta}+\tilde{a}))]\\ &= \tilde{\theta}+\Phi[\sqrt{\alpha}(\underline{x}-(\theta_{\underline{x}}+\tilde{a}))]-\Phi[\sqrt{\alpha}(\underline{x}-(\tilde{\theta}+\tilde{a}))]\\ &\leq \tilde{\theta}\end{aligned}$$

where the last inequality follows because the hypothesis $\tilde{\theta}\leq\theta_{\underline{x}}$ implies $\Phi[\sqrt{\alpha}(\underline{x}-(\tilde{\theta}+\tilde{a}))]\geq\Phi[\sqrt{\alpha}(\underline{x}-(\theta_{\underline{x}}+\tilde{a}))]$. This is a contradiction, and so $\tilde{\theta}>\theta_{\underline{x}}$. Therefore, the regime is overthrown for at least all $\theta<\theta_{\underline{x}}$. A parallel argument shows that if it is strictly dominant for $s(x)=0$ for all $x>\overline{x}$, then the regime is not overthrown for at least all $\theta>\theta_{\overline{x}}$ where $\theta_{\overline{x}}$ is defined analogously. □

Proof of Lemma 6. Since $K(\hat{x},x)$ is continuously differentiable in x, an application of the implicit function theorem to (21) shows that ψ is continuous. Fixed points of ψ satisfy $x^*=\psi(x^*)$. Equivalently, by Lemma 3, they satisfy $K(x^*,x^*)=P(\theta(x^*)-x^*,0,0)=p$, where, in another slight abuse of notation, $\theta(\hat{x})$ is the critical state in the regime's problem (13)-(14). By Lemma 2 and the intermediate value theorem there is a unique $z^*\in\mathbb{R}$ such that $P(z^*,0,0)=p$. Applying the implicit function theorem to (13)-(14) gives

$$\theta'(\hat{x})=\frac{\sqrt{\alpha}\phi\{\sqrt{\alpha}[\hat{x}-\theta(\hat{x})-a_{\hat{x}}(\theta(\hat{x}))]\}}{1+\sqrt{\alpha}\phi\{\sqrt{\alpha}[\hat{x}-\theta(\hat{x})-a_{\hat{x}}(\theta(\hat{x}))]\}}\in(0,1) \tag{45}$$

Since $\theta(-\infty)=0$ and $\theta(\infty)=1$, there is a unique $x^*\in\mathbb{R}$ such that $\theta(x^*)-x^*=z^*$, hence ψ has a unique fixed point, the same x^*. Now using Lemma 3 and implicitly differentiating (21) we have

$$\psi'(\hat{x})=1+\frac{P_\theta[\theta(\hat{x})-\hat{x},\psi(\hat{x})-\hat{x},0]}{P_x[\theta(\hat{x})-\hat{x},\psi(\hat{x})-\hat{x},0]}[1-\theta'(\hat{x})] \tag{46}$$

By Lemma 2, $P_\theta>0$ and $P_x<0$ and $\theta'(\hat{x})\in(0,1)$ from (45). Therefore $\psi'(\hat{x})<1$ for all $\hat{x}$. To see that $\psi'(x^*)>0$, first notice that it is sufficient that $P_\theta/P_x\geq -1$ when evaluated at $\hat{x}=x^*$. Calculating the derivatives shows that this is true if and only if

$$\phi[\sqrt{\alpha}(y(\theta^*)-x^*)]+\int_{\theta^*}^{\infty}\sqrt{\alpha}\phi'[\sqrt{\alpha}(y(t)-x^*)]dt\leq 0 \tag{47}$$

where $\theta^*:=\theta(x^*)$ and where $y(t)=t+a(t)$ is the mean of the signal distribution from which a citizen is sampling if the regime has type $t\geq\theta^*$. To show that this condition always holds, we need to consider two separate cases: (i) where costs are linear, and (ii) where costs are strictly convex. If costs are linear, $C(a)=ca$, then if $c\geq\bar{c}:=\sqrt{\alpha}\phi(0)$ the result is trivial because $a(t)=0$ for all $t\in\mathbb{R}$. So suppose $c<\bar{c}$. Then $a(t)=\max[0,x^*+\gamma-t]$ where $\gamma:=\sqrt{2\log\left(\sqrt{\alpha}\phi(0)/c\right)/\alpha}>0$. Calculating the integral and then simplifying shows that (47) holds if and only if $-\alpha\gamma\phi(\sqrt{\alpha}\gamma)a(\theta^*)\leq 0$ which is true because $a(\theta^*)\geq 0$. Now consider case (ii) where costs are strictly convex. From the optimality conditions for the regime's choice of action we have that $a(t)>0$ for all $t\geq\theta^*$ and

$$\sqrt{\alpha}\phi[\sqrt{\alpha}(y(t)-x^*)]=C'[a(t)],\quad t\geq\theta^* \tag{48}$$

Differentiating with respect to t gives

$$\alpha\phi'[\sqrt{\alpha}(y(t)-x^*)]y'(t)=C''[a(t)]a'(t),\quad t\geq\theta^* \tag{49}$$

Using the associated second order condition shows that $y'(t) > 0$ for $t \geq \theta^*$. Since y is invertible, a change of variables shows that (47) holds if and only if

$$\int_{\theta^*}^{\infty} \phi'[\sqrt{\alpha}(y(t) - x^*)] \frac{a'(t)}{y'(t)} dt \geq 0 \tag{50}$$

Using (49) we equivalently have the condition

$$\int_{\theta^*}^{\infty} \frac{\phi'[\sqrt{\alpha}(y(t) - x^*)]^2}{C''[a(t)]} dt \geq 0 \tag{51}$$

which is true since the integrand is non-negative. Therefore, $P_\theta / P_x \geq -1$ at $\hat{x} = x^*$ and $\psi'(x^*) > 0$.

Finally, $\psi(\hat{x}) \leq x^*$ for every $\hat{x} < x^*$ is proven by contradiction. Suppose not. Then by continuity of ψ there exists $\tilde{x} < x^*$ such that $\psi(\tilde{x}) = x^*$. Moreover, since $\psi'(x^*) > 0$, we must have $\psi'(\tilde{x}) < 0$ for at least one such $\tilde{x}$. Since $\psi(\tilde{x}) = x^*$ and $K(x^*, x^*) = p$, under this hypothesis we can write $K[\psi(\tilde{x}), \psi(\tilde{x})] = p$ so by the implicit function theorem $\psi(\tilde{x})$ must satisfy

$$\psi'(\tilde{x})[K_1(x^*, x^*) + K_2(x^*, x^*)] = 0 \tag{52}$$

where the hypothesis $\psi(\tilde{x}) = x^*$ is used to evaluate the derivatives of K. Since $\psi'(\tilde{x}) < 0$, this can only be satisfied if $K_1(x^*, x^*) + K_2(x^*, x^*) = 0$. But for any $\hat{x} \in \mathbb{R}$, the value $\psi(\hat{x})$ is implicitly defined by $K[\hat{x}, \psi(\hat{x})] = p$ so that by the implicit function theorem $\psi'(\hat{x}) = -K_1[\hat{x}, \psi(\hat{x})]/K_2[\hat{x}, \psi(\hat{x})]$. From (46) we know $\psi'(\hat{x}) < 1$ for any $\hat{x}$ and since $K_2 < 0$ from Lemma 3 we conclude $K_1[\hat{x}, \psi(\hat{x})] + K_2[\hat{x}, \psi(\hat{x})] < 0$ for any $\hat{x}$. For $\hat{x} = x^*$ in particular, $K_1(x^*, x^*) + K_2(x^*, x^*) < 0$ so we have the needed contradiction. Therefore $\psi(\hat{x}) \leq x^*$ for every $\hat{x} < x^*$. A symmetric argument shows $\psi(\hat{x}) \geq x^*$ for every $\hat{x} > x^*$. □

Proof of Theorem 1. There is a unique equilibrium in monotone strategies. To show this, we take arbitrary $\hat{x} \in \mathbb{R}$ and solve the regime's problem to get $\theta_{\hat{x}}$ and $a_{\hat{x}}(\theta) = a(\theta - \hat{x})$ from Lemma 1. We use these functions to construct $K(\hat{x}, x_i)$ for each signal $x_i \in \mathbb{R}$ using (16) and use Lemma 3 to conclude that $K(\hat{x}, \hat{x}) = P(\theta_{\hat{x}} - \hat{x}, 0, 0)$. Then let $\mu(z) := P(z, 0, 0)$ and use Lemma 3 and the intermediate value theorem to deduce that there is a unique $z^* \in \mathbb{R}$ such that $\mu(z^*) = p$. This gives a unique difference $z^* = \theta^* - x^*$ that can be plugged into the regime's indifference condition (20) to get the unique $\theta^* \in [0, 1)$ such that the regime is overthrown if and only if $\theta < \theta^*$. The unique signal threshold is then $x^* = \theta^* - z^*$ and the unique hidden action function is given by, in a slight abuse of notation, $a(\theta) := a(\theta - x^*)$.

To show there is no other equilibrium, we use Lemma 4 to establish *dominance regions* delineated by $\underline{x}, \overline{x}$, both finite, such that it is an (iteratively) dominant strategy for $s(x) = 1$ for all $x < \underline{x}$ and for $s(x) = 0$ for all $x > \overline{x}$. This leaves open the behavior of $s(x)$ for $x \in [\underline{x}, \overline{x}]$. We then use Lemma 5 to conclude that if it is strictly dominant for $s(x) = 1$ for all $x < \underline{x}$, then the regime is overthrown for at least all $\theta < \theta_{\underline{x}}$ where $\theta_{\underline{x}} > 0$ and solves (13)-(14). Since cumulative distribution functions are non-decreasing, for any beliefs of the citizens, the posterior probability assigned by a citizen to the regime's overthrow is *at least* as much as the probability they assign to $\theta < \theta_{\underline{x}}$. Therefore the expected gain from $s(x) = 1$ to a citizen with signal x is at least $K(\underline{x}, x_i) - p$. Lemma 3 and the intermediate value theorem imply there is a unique $\psi(\underline{x})$ solving $K[\underline{x}, \psi(\underline{x})] = p$ such that if it is strictly dominant for $s(x) = 1$ for all $x < \underline{x}$, then it is strictly dominant for $s(x) = 1$ for *at least* all $x < \psi(\underline{x})$. A parallel argument establishes that if is it strictly dominant for $s(x) = 0$ for all $x > \overline{x}$, then it is strictly dominant for $s(x) = 0$ for at least all $x > \psi(\overline{x})$.

Now let $\underline{x}_0 := \underline{x}$ and $\overline{x}_0 := \overline{x}$ and generate sequences $\{\underline{x}_n\}_{n=0}^{\infty}$ from $\underline{x}_{n+1} = \psi(\underline{x}_n)$ and $\{\overline{x}_n\}_{n=0}^{\infty}$ from $\overline{x}_{n+1} = \psi(\overline{x}_n)$. By Lemma 6 ψ is continuous, has a unique fixed point $x^* = \psi(x^*)$ (the same

x^* as above) with $\psi'(x^*) \in (0,1)$ and $\psi(\underline{x}_n) \leq x^*$ for all $\underline{x}_n < x^*$. Therefore $\{\underline{x}_n\}_{n=0}^{\infty}$ is bounded above, strictly monotone increasing and so converges $\underline{x}_n \nearrow x^*$ as $n \to \infty$. Similarly $\{\overline{x}_n\}_{n=0}^{\infty}$ is bounded below, strictly monotone decreasing and so converges $\overline{x}_n \searrow x^*$. After n iterations, the only candidates for an equilibrium strategy all have $s(x) = 1$ for $x < \underline{x}_n$ and $s(x) = 0$ for $x > \overline{x}_n$ with $s(x)$ arbitrary for $x \in [\underline{x}_n, \overline{x}_n]$. In the limit as $n \to \infty$, the only strategy that survives the elimination of strictly dominated strategies is the one with $s(x) = 1$ for $x < x^*$ and $s(x) = 0$ otherwise. Therefore the only equilibrium is the unique monotone equilibrium. □

Proof of Theorem 2. For each precision α, there is a unique equilibrium. I find a unique solution to a constrained problem consisting of the original system of nonlinear equations plus a set of constraints that govern the asymptotic behavior of the endogenous variables. But, because the equilibrium conditions have a unique solution for each α, the solution to the original problem and to the constrained problem coincide.

The equilibrium conditions can be written

$$(1-p)\Phi[\sqrt{\alpha}\,(\theta_\alpha^* - x_\alpha^*)] = p\int_{\theta_\alpha^*}^{\infty} \sqrt{\alpha}\phi[\sqrt{\alpha}(x_\alpha^* - \theta - a_\alpha(\theta - x_\alpha^*))]d\theta \tag{53}$$

and

$$\theta_\alpha^* = \Phi[\sqrt{\alpha}(x_\alpha^* - \theta_\alpha^* - a_\alpha(\theta_\alpha^* - x_\alpha^*))] + C[a_\alpha(\theta_\alpha^* - x_\alpha^*)] \tag{54}$$

with actions characterized by

$$C'[a_\alpha(\theta - x_\alpha^*)] = \sqrt{\alpha}\phi[\sqrt{\alpha}(x_\alpha^* - \theta - a_\alpha(\theta - x_\alpha^*))], \qquad \theta \geq \theta_\alpha^* \tag{55}$$

Now let $\alpha \to \infty$. The auxiliary constraints that govern the asymptotic behavior of the endogenous variables are assumed to be

$$\lim_{\alpha\to\infty} \sqrt{\alpha}(x_\alpha^* - \theta_\alpha^* - a_\alpha(\theta_\alpha^* - x_\alpha^*)) = \lim_{\alpha\to\infty} \sqrt{\alpha}(\theta_\alpha^* - x_\alpha^*) = -\infty \tag{56}$$

If condition (56) holds, from (54) we have $\theta_\alpha^* = C[a_\alpha(\theta_\alpha^* - x_\alpha^*)]$. Similarly, if (56) holds, then $\Phi[\sqrt{\alpha}\,(\theta_\alpha^* - x_\alpha^*)] \to 0$ and the value of the integral on the right hand side of (53) converges to zero. From (53) and (55), this requires

$$\lim_{\alpha\to\infty} \int_{\theta_\alpha^*}^{\infty} C'[a_\alpha(\theta - x_\alpha^*)]d\theta = 0$$

Since $\theta_\alpha^* \in [0,1]$ and $C'[a_\alpha(\theta - x_\alpha^*)] \geq 0$ and is uniformly continuous in α, this can only be true if $a_\alpha(\theta - x_\alpha^*) \to 0^+$ for all $\theta \geq \theta_\alpha^*$. But then if $a_\alpha(\theta_\alpha^* - x_\alpha^*) \to 0^+$, $C[a_\alpha(\theta_\alpha^* - x_\alpha^*)] \to 0^+$ and so $\theta_\alpha^* \to 0^+$ too. Finally, if both constraints are to hold simultaneously for large α, $x_\alpha^* - \theta_\alpha^*$ is positive and $x_\alpha^* - \theta_\alpha^* - a_\alpha(\theta_\alpha^* - x_\alpha^*)$ is negative. For both constraints to have the same sign, x_α^* can neither diverge nor converge to either a strictly positive or a strictly negative number. So $x_\alpha^* \to 0^+$. Hence we have found a solution to the constrained problem.

Now for the second part of the Theorem. Recall that for this part we assume strictly convex costs. Let $\alpha \to 0^+$ such that $\sqrt{\alpha}x_\alpha^* \to \infty$ holds. Then $x_\alpha^* \to \infty$. Since $\theta_\alpha^* \in [0,1]$, we have $\sqrt{\alpha}\,(x_\alpha^* - \theta_\alpha^*) \to \infty$ and the integral on the right hand side of (53) must converge to zero. Hence, by (55), $a_\alpha(\theta - x_\alpha^*) \to 0^+$ for all $\theta \geq \theta_\alpha^*$ (the strict convexity of C is assumed here so that (55) holds for all θ even as $\alpha \to 0^+$; with constant marginal costs, this would not be true). But if $a_\alpha(\theta_\alpha^* - x_\alpha^*) \to 0^+$, $\theta_\alpha^* \in [0,1]$, and $\sqrt{\alpha}x_\alpha^* \to \infty$, then (54) requires that $\theta_\alpha^* \to 1^-$. Once again we have found a solution to the constrained problem. □

B The role of coordination: regime vs. single agent

This appendix highlights the role of imperfect coordination in enabling the regime to survive even when signals are precise. Suppose to the contrary that citizens are perfectly coordinated and receive one $x := \theta + a + \varepsilon$. Collectively, they can overthrow the regime if $\theta < 1$. In a monotone equilibrium the mass attacks the regime, $S(x) = 1$, if and only if $x < x^*$ where x^* solves $\Pr(\theta < 1|x^*) = p$.

The regime now faces *aggregate uncertainty*. It does not know what value of x will realize. The regime chooses its hidden action to maximize its expected payoff

$$\int_{-\infty}^{\infty} \max[0, \theta - S(x)] f(x|\theta, a) dx - C(a) = -\min[\theta, 1]\Phi[\sqrt{\alpha}(x^* - \theta - a)] - C(a) \tag{57}$$

Regimes with $\theta < 0$ are overthrown and so never engage in costly manipulation.

Example: strictly convex costs. Suppose, with some loss of generality, that costs are *strictly convex*, $C''(a) > 0$. This implies all regimes $\theta > 0$ will choose some positive manipulation $a(\theta) > 0$ even regimes that are overthrown ex post. The key first order necessary condition for the regime's choice of action $a(\theta)$ is

$$\min[\theta, 1]\sqrt{\alpha}\phi[\sqrt{\alpha}(x^* - \theta - a) = C'(a), \quad \theta \geq 0 \tag{58}$$

As usual, there may be two solutions to this first order condition: if so, the smaller is eliminated by the second order condition. An equilibrium of this game is constructed by simultaneously determining $a(\theta)$ and the x^* that solves $\Pr(\theta < 1|x^*) = p$.

The first order condition implies that taking as given x^* the regime's $a(\theta) \to 0^+$ as $\alpha \to \infty$. Given this, the probability of overthrowing the regime $\Pr(\theta < 1|x) \to \mathbb{1}\{x < 1\}$ as $\alpha \to \infty$. This implies $x^* \to 1$. With arbitrarily precise information, the regime takes no action and so x is very close to θ. The mass attacks only if it believes $\theta < 1$ and since x is close to θ attacks only if $x < 1$. So if citizens are perfectly coordinated then for precise information regime change occurs for all $\theta < 1$. By contrast, Theorem 2 tells us that if citizens are imperfectly coordinated then for precise information all regimes $\theta \geq 0$ survive.

Angeletos, Hellwig, and Pavan (2006) provide a related analysis. In their model, if agents are imperfectly coordinated then for precise information θ^* can be any $\theta \in (0, \theta^*_{\text{MS}}]$ where $\theta^*_{\text{MS}} = 1 - p < 1$. But if agents are perfectly coordinated then for precise information regime change occurs for all $\theta < 1$. Thus when information is precise the two models agree about the regime change outcome when agents are perfectly coordinated but come to different conclusions when agents are imperfectly coordinated.

References

Angeletos, G.-M., C. Hellwig, and A. Pavan (2006): "Signaling in a global game: Coordination and policy traps," *Journal of Political Economy*, 114(3), 452–484. 1, 3, 5, 12, 32

——— (2007): "Dynamic global games of regime change: Learning, multiplicity, and the timing of attacks," *Econometrica*, 75(3), 711–756. 1

Angeletos, G.-M., and I. Werning (2006): "Crises and prices: Information aggregation, multiplicity and volatility," *American Economic Review*, 96(5), 1720–1736. 1, 23

Atkeson, A. (2000): "Discussion," in *NBER Macroeconomics Annual*, ed. by B. S. Bernanke, and K. Rogoff, pp. 162–171. MIT Press. 2

Bannier, C. E., and F. Heinemann (2005): "Optimal transparency and risk-taking to avoid currency crises," *Journal of Institutional and Theoretical Economics*, 161, 374–391. 25

Carlsson, H., and E. van Damme (1993): "Global games and equilibrium selection," *Econometrica*, 61(5), 989–1018. 2, 7

Chamley, C. P. (1999): "Coordinating regime switches," *Quarterly Journal of Economics*, 114(3), 869–905. 2

——— (2004): *Rational herds: Economic models of social learning*. Cambridge University Press. 2

Cheli, B., and P. Della-Posta (2002): "Self-fulfilling currency attacks with biased signals," Universita di Pisa. 2

Cooper, R. W. (1999): *Coordination games: Complementarities and macroeconomics*. Cambridge University Press. 2

Crawford, V. P., and J. Sobel (1982): "Strategic information transmission," *Econometrica*, 50(6), 1431–1451. 3, 12

Frommer, B. (2005): *National cleansing: Retribution against Nazi collaborators in postwar Czechoslovakia*. Cambridge University Press. 6

Heinemann, F., and G. Illing (2002): "Speculative attacks: Unique equilibrium and transparency," *Journal of International Economics*, 58, 429–450. 25

Hellwig, C. (2002): "Public information, private information, and the multiplicity of equilibria in coordination games," *Journal of Economic Theory*, 107, 191–222. 7, 23

Jackson, J. (2001): *France: The dark years, 1940-1944*. Oxford University Press. 6

Kalathil, S., and T. C. Boas (2003): *Open networks, closed regimes: The impact of the internet on authoritarian rule*. Carnegie Endowment for International Peace. 2

Matthews, S. A., and L. J. Mirman (1983): "Equilibrium limit pricing: The effects of private information and stochastic demand," *Econometrica*, 51(4), 981–996. 3, 12

Metz, C. (2002): "Private and public information in self-fulfilling currency crises," *Zeitschrift fur Nationalokonomie*, 76(1), 65–85. 23, 24

Milgrom, P., and J. Roberts (1982): "Limit pricing and entry under incomplete information: An equilibrium analysis," *Econometrica*, 50(2), 443–460. 12

Morris, S., and H. S. Shin (1998): "Unique equilibrium in a model of self-fulfilling currency attacks," *American Economic Review*, 88(3), 587–597. 2, 7

——— (2000): "Rethinking multiple equilibria in macroeconomic modeling," in *NBER Macroeconomics Annual*, ed. by B. S. Bernanke, and K. Rogoff, pp. 139–161. MIT Press. 7, 23

——— (2003): "Global games: Theory and applications," in *Advances in Economics and Econometrics: Theory and Applications*, ed. by M. Dewatripont, L. P. Hansen, and S. J. Turnovsky. Cambridge University Press. 2, 7, 23

——— (2004): "Coordination risk and the price of debt," *European Economic Review*, 48(1), 133–153. 2, 7, 23

Obstfeld, M. (1986): "Rational and self-fulfilling balance-of-payments crises," *American Economic Review*, 76(1), 72–81. 2, 6

——— (1996): "Models of currency crises with self-fulfilling features," *European Economic Review*, 40(3–5), 1037–1047. 2, 6

Riley, J. G. (2001): "Silver signals: Twenty-five years of screening and signaling," *Journal of Economic Literature*, 39(2), 432–478. 12

Tirole, J. (1988): *The theory of industrial organization*. MIT Press. 12

Wilson, R. (1992): "Strategic models of entry deterrence," in *Handbook of game theory*, ed. by R. J. Aumann, and S. Hart. North-Holland/Elsevier Science. 12

Zeman, Z. (1973): *Nazi propaganda*. Oxford University Press, second edn. 18