Beeps*

Jeffrey C. Ely †

January 28, 2016

Abstract

I introduce and study dynamic persuasion mechanisms. A principal privately observes the evolution of a stochastic process and sends messages over time to an agent. The agent takes actions in each period based on her beliefs about the state of the process and the principal wishes to influence the agent's action. I characterize the optimal persuasion mechanism and show how to derive it in applications. I then consider the extension to multiple agents where higher-order beliefs matter.

Keywords: beeps, information design, obfuscation principle.

*This research is supported through grants from the National Science Foundation. Thanks especially to Toomas Hinnosaar and Kevin Bryan for early discussions on this topic. Four anonymous referees and the editor gave valuable suggestions which have improved the paper and for which I am grateful. I received excellent questions, comments, and suggestions from Gabriel Carroll, Jerome Renault, Johannes Horner, Moritz Meyer-ter-Vehn, Simon Board, Sergiu Hart, Andrea Prat, Tanjim Hossain, Martin Szydlowski, Eilon Solan, Nicolas Vielle, Debraj Ray, Daria Khromenkova and Stephen Morris. This paper began as a March 2012 blog post at cheaptalk.org.

†Charles E. and Emma H. Morrison Professor of Economics, Northwestern University. jeff@jeffely.com

1 Introduction

In long-run relationships the control of information is an important instrument for coordinating and incentivizing actions. In this paper I analyze the optimal way to filter the information available to an agent over time in order to influence the evolution of her beliefs and therefore her sequence of actions.

For example, the seller of an object which is depreciating stochastically over time must decide what information to disclose to potential buyers about the current quality. A supervisor must schedule performance evaluations for an agent who is motivated by career concerns.[1] A planner may worry that self-interested agents experiment too little or herd too much, and can use filtered information about the output of experiments to control the agent's motivations.[2]

The common theme in all such applications is that messages that motivate the agent must necessarily be coupled with messages that harm future incentives and the principal controls the timing of these messages. For example, if the seller can credibly signal that the depreciation has been slow, then in the absence of such a signal the buyers infer that the object has significantly decreased in value. The dynamic context adds a time dimension to this tradeoff: information revealed today not only affects current incentives but also alters the effectiveness of information revealed tomorrow. Optimal information management designs the magnitude as well as the timing of disclosures.

I develop a general model to analyze dynamic information disclosure. Formally the model is a dynamic extension of the Bayesian Persuasion model of Kamenica and Gentzkow (2011). A principal privately observes the evolution of a stochastic process and sends messages over time to an agent. The agent is myopic[3], taking actions in each

[1] For a related model see Orlov (2013).

[2] Two recent papers studying related problems are Che and Hörner (2013) and Kremer, Mansour and Perry (2013).

[3] This is without loss of generality if the agent's action is unobservable and/or non-contractible. See

period based on her beliefs about the state of the process and the principal wishes to influence the agent's action. Relative to the static model of Kamenica and Gentzkow (2011), dynamics add several interesting dimensions to the incentive problem. The state is evolving so even if the principal offers no independent information, the agent's beliefs will evolve autonomously. Information management can shape the path of this evolution. Messages that persuade the agent to take desired actions today also alter the path of beliefs in the future. There is a tradeoff between current persuasion and the ability to persuade in the future.

To illustrate these ideas consider the following example. The IT department in an organization is concerned about workers spending too much time reading email and is considering how to configure workstations to minimize distractions. Currently the enterprise email software is configured to beep when an email arrives giving workers an irresistible temptation to check their email and inevitably succumb to further distractions such as web surfing, social media etc.

The IT department considers turning the email beep off to avoid the immediate temptation it induces.[4] However there is no free lunch: knowing that the beep is turned off, the worker knows that it is becoming increasingly likely as time passes that an email has (silently) arrived. Eventually it will be sufficiently likely that she will again give in to temptation and check. Let us suppose that her temptation is described by a threshold p^* such that she will stop working as soon as she believes with probability greater than p^* that an email is waiting to be read.

We can calculate the productivity trade-off between beep-on and beep-off. If λ is the Poisson arrival rate of email, then $1/\lambda$ is the expected length of time before an email arrives. With the beep turned on, $1/\lambda$ is also the expected length of time spent working

the concluding section for further discussion.

[4]Completely disabling email or the web is not an option as these are essential to the organization and distinguishing between productive and unproductive uses of these facilities is prohibitively costly.

before the beep sounds and the worker is distracted. By contrast when the beep is turned off the duration she spends working is deterministic and is given by

$$t^* = -\frac{\log(1-p^*)}{\lambda},$$

because t^* is how long it takes for the probability of an arrival $1 - e^{-\lambda t}$ to pass the threshold p^*. Comparing these expressions we see that beep-on yields a longer expected waiting time when p^* is low but that beep-off is better when p^* is high[5] but as I will show the optimal mechanism outperforms either of these.

In terms of the model in this paper the IT department is the principal and the worker is the agent. There is an uncertain state of the world (whether email is waiting) that is stochastically changing over time and observed only by the principal. The problem for the principal is how to send filtered information to the agent about the state in order to influence the agent's behavior over time, here whether to work or shirk. Among all feasible policies for managing the agent's information, beep-on and beep-off are the two extremes: perfect information and no information respectively.

Between these there is an enormous range of intermediate polices. For example, consider a random beep: when an email arrives the software performs a randomization and with probability $z \in (0,1)$ emits a beep. Relative to beep-on (which is just $z = 1$) this slows the arrival rate of beeps, and relative to beep-off ($z = 0$) it slows the rate at which beliefs increase in the absence of a beep. An interior z typically outperforms either extreme.[6] And in addition to randomization, the principal can use time and history-dependent policies, and even messages that vary along a continuum (the

[5]Interestingly, when the worker is easily distracted (represented by a low p^*) IT should nevertheless amplify distractions by turning the beep on. This is because in return for the distraction, she is able to remain working as long as email has yet to arrive. This can lead to a longer waiting time on average. On the other hand it is best to disable the beep for a worker with a high p^*. The precise turning point is the threshold that satisfies $1 = -\log(1-p^*)$ which is $1 - 1/e$, roughly 0.63.

[6]In an online appendix I work through detailed calculations of these mechanisms and others.

"volume" of the beep) to further fine-tune the evolution of the agent's beliefs.

Characterizing the full set of these dynamic information policies for the purposes of optimization is complex. In this paper, by extending the Kamenica and Gentzkow (2011) and Aumann and Maschler (1995) tools of static persuasion problems, I develop a tractable characterization of dynamic information management policies and a method for optimizing over these in a wide class of problems including many states, many actions, and a general stochastic process.

When I apply these techniques to the special case of the email example (two actions, two states, one of which is absorbing) I find a simple qualitative prescription for information management. First, the principal strictly prefers not to interfere during times when the agent is already, even temporarily, willing to act in the principal's interest. The principal should simply stand by and allow the agent's beliefs to drift toward the point where incentives become misaligned. Intuitively, any intervention that moves the agent farther from the threshold must with positive probability also move the agent closer. In the dynamic setting, the principal is endogenously risk averse over gambles that add randomness to the duration the agent spends working.[7] However once that threshold (p^* in the example) is reached the optimal mechanism begins randomly conceding: with strictly positive probability the principal informs the agent that the state has changed and the agent stops working. The stochastic rate of these concessions is chosen to be as small as possible subject to the following incentive constraint. Conditional on *no* such signal, the agent infers that the probability of email is exactly at the threshold and he goes on working. In particular, good behavior is completely front-loaded: along every path of the process the agent changes behavior exactly once, from working to shirking.[8]

[7] Formally the value function arising out of the principal's optimization is strictly concave.

[8] This frontloading is not due to discounting but rather a further manifestation of the principal's endogenous risk aversion. Any mechanism which induces the agent to return to work necessarily adds additional randomness.

Indeed, I show that the mechanism that achieves this optimum is a beep with a *delay.* In particular, if email arrives at date t the beep sounds at date $t + t^*$. This mechanism keeps the worker working for an expected duration $t^* + 1/\lambda$, i.e. the *sum* of the durations from the two extreme policies. Here I will show how the mechanism achieves this, and in Section 2, as an illustration of the general techniques in the paper I will show that this is the best the principal can do.[9]

With a delay of length t^* the agent knows he will not be informed of any arrival prior to date t^*. Thus, the agent is effectively facing a beep-off policy prior to that date and therefore his beliefs trend upward until they reach the threshold p^*. The agent works throughout this time and his accounts for the first term in the expression above.

Consider now any date after t^*. If the beep sounds the worker learns that an email has arrived (t^* moments ago) and she stops working. But suppose instead that no beep is heard. The worker infers that no email has arrived at any time earlier than t^* moments ago, but she learns nothing more than that. By construction, p^* is the conditional probability of an email arrival given no information for a timespan of length t^*. Thus the worker's beliefs remain fixed at p^*, and she continues to work until she hears a beep. The fixed delay means that the arrival rate of beeps is the same as the arrival rate of email, i.e. λ. Therefore, beginning at date t^*, the agent continues to work for an additional $1/\lambda$ duration (in expectation), accounting for the second term.

The email beeps problem is the simplest example in that there are two states (one absorbing), two actions for the agent, and the principal's payoff is state-independent. Following a detailed analysis of the example I turn to the general problem with an arbitrary finite state space, and a general payoff function which accommodates an arbitrary action space for the agent and possibly state-dependent preferences for the principal. I turn then to the somewhat more elaborate example of a monopolist providing informa-

[9]This implementation was first conjectured by Toomas Hinnosaar.

tion about the quality of a good that depreciates over time. When demand is sensitive to quality a full-disclosure policy is optimal, whereas when demand is less sensitive to quality the optimal policy is uninformative. In an intermediate case when demand pools consumers of both types, the optimal policy is partially informative, initially providing imperfect signals of quality but eventually reverting to non-disclosure.

Finally I consider the case of multiple agents who interact strategically. In this case, to control the incentives of each agent the principal must not only manage beliefs about the state, but also beliefs about the actions of other agents. The setting is a stylized model of a bank run. The bank can release public and private information to the agents over time in order to prevent them from coordinating a run on deposits. I analyze this problem and show how the bank optimally uses private and minimally correlated disclosures to achieve this.

1.1 Related Literature

The model studied in this paper is a dynamic extension of the static Bayesian persuasion model of Kamenica and Gentzkow (2011). It adapts results from that paper to characterize the optimal value function and optimal policy in a dynamic mechanism design problem in which the principal controls the flow of information available to the agent. Two key methods from Kamenica and Gentzkow (2011), a tractable characterization of feasible policies and a geometric characterization of the optimum, were in turn adapted from the study of repeated games with incomplete information due to Aumann and Maschler (1995).

Concurrently and independently of this work, Renault, Solan and Vieille (2014) examined the baseline single-agent two-action model in this paper but with a different focus. Renault, Solan and Vieille (2014) are interested in cases in which the optimal policy is "greedy," that is in each period the principal acts as if he is maximizing today's

payoff without regard to future consequences.[10]

Static Bayesian persuasion models have multi-agent counterparts in the literature. Bergemann and Morris (2013) develop a concept of Bayes Correlated Equilibrium and show that the set of all such solutions is equal to the set of all action distributions that can be induced by a principal who can control the information structure of the agents. These results are applied to static games with strategic substitutes and complements in Bergemann and Morris (2014) and to the analysis of information structures in auctions in Bergemann, Brooks and Morris (2014). Taneva (2014) presents a complete analysis of two-state, two-player, two-action static games.

There is a large literature on the design of information feedback in dynamic principal-agent problems and games. Che and Hörner (2013) and Kremer, Mansour and Perry (2013) both study information feedback in a social experimentation problem in which agents decide in sequence whether to experiment with a new technology. Halac, Kartik and Liu (2014) consider the optimal combination of monetary incentives and information revelation policies in a research contest. Gershkov and Szentes (2009) study information aggregation in organizations in which information acquisition is private and costly. Lizzeri, Meyer and Persico (2002) study interim performance evaluations as an incentive device. Aoyagi (2010) studies the effects of information feedback in a dynamic tournament. Angeletos, Hellwig and Pavan (2007) and Morris and Shin (1999) study learning about fundamentals in the context of dynamic coordination games and currency crises, respectively. These papers generally consider exogenous information structures or compare a few extreme policies (full disclosure, public disclosure, no disclosure). The methods in this paper may help push the literature forward by optimizing over the full range of information policies.[11]

[10]The monopoly example I analyze in Section 4 is a case of a non-greedy optimal policy, the online appendix has another example.

[11]Gershkov and Szentes (2009) is one notable exception where the optimal information structure is characterized.

2 Basic Model

This section introduces the main ideas of the paper through an analysis of the basic example from the introduction. In this basic model there are two states $s \in S = \{0, 1\}$ and the agent has two actions, $a \in \{0, 1\}$. The process begins in state $s = 0$ and transitions to state $s = 1$ at Poisson rate $\lambda > 0$. State $s = 1$ is absorbing.

The agent forms beliefs μ_t at each date t about the current state. Specifically μ_t is the probability that the agent assigns to state $s = 1$. The agent's only source of information about the current state is the principal.[12] Because the agent knows that transitions happen at rate λ, in the absence of any additional information from the principal her beliefs evolve over time as follows

$$\frac{d\mu}{dt} = \lambda(1 - \mu_t). \tag{1}$$

The principal can additionally influence the evolution of beliefs through his information policy as discussed below.

The agent wishes to match his action to the state. Specifically, she strictly prefers action $a = 1$ when $\mu > p^*$, she strictly prefers action $a = 0$ when $\mu < p^*$, and she is indifferent whenever $\mu = p^*$. On the other hand, the principal has state-independent preferences and always strictly prefers the agent to choose $a = 0$. We will normalize the principal's flow payoffs so that he earns 1 whenever the agent chooses action $a = 0$ and zero otherwise. Because the agent's action will be determined by her belief μ_t, we

[12] In the Conclusion I discuss how the model can accommodate external information that is *public* and discuss the new issues that arise with *private* information acquisition by the agent.

can collapse the principal's flow payoff to a function[13] that depends directly on μ.

$$u(\mu_t) = \begin{cases} 1 & \text{if } \mu_t \leq p^* \\ 0 & \text{if } \mu_t > p^* \end{cases}$$

The principal chooses an information policy to maximize his expected discounted long-run payoff, where $r \in (0, \infty)$ is the discount rate.

The set of feasible information policies is completely general: there is an arbitrary set of messages M and a rule $\sigma_t(h)$ which specifies for each date t a probability distribution over messages as a function of the past history. As discussed in the introduction, such policies include the extremes of full information, no information, and all conceivable intermediate policies. The full-information policy, or beep-on, is the rule that reveals perfectly the current state s_t: i.e., $\sigma_t(s_t) = s_t$. A policy of no information, beep-off, is obtained when the set of messages M is a singleton.

Any policy determines the stochastic evolution of the agent's beliefs. For example, faced with the full-information policy the agent's belief at each point in time will be either $\mu_t = 1$ (knowing for sure that the state is $s = 1$) or $\mu_t = 0$ (knowing for sure that the state is $s = 0$). Under the no-information policy the agent's beliefs will equal $\mu_t = (1 - e^{-\lambda t})$, the unconditional probability that an email has arrived before date t.

Policies can be designed to fine-tune the agent's beliefs even further. For example, there exists a policy in which the principal randomizes over messages in the interval $[0, 1]$ (the "volume" of the beep) in such a way that when the agent receives message m he infers that the probability of state $s = 1$ is exactly m.[14]

[13] In doing so we are implicitly "breaking ties" at $\mu = p^*$ in favor of the principal, a standard technical convenience in principal-agent models.

[14] Message m is sent with probability density

$$f_1(m) = \frac{(\alpha+1)m^{\alpha+1}}{\mu_t} \quad \text{and} \quad f_0(m) = \frac{(\alpha+1)m^{\alpha}(1-m)}{1-\mu_t}$$

In such a policy the principal is essentially telling the agent what his belief should be and because of the way the probabilities have been chosen, the agent rationally accepts the suggestion. By analogy to traditional mechanism design, I call these *direct obfuscation mechanisms.* Indeed, analogously to the revelation principle, previous authors including Aumann and Maschler (1995), have shown in a static context it is without loss of generality to focus on direct mechanisms: any distribution over induced beliefs can be generated by such a mechanism provided that distribution satisfies the law of total probability, i.e. that the agent's expectation of the induced belief equals his prior.

In the static persuasion setting of Kamenica and Gentzkow (2011) this *obfuscation principle* allows us to cut through the vast set of information policies focus on direct mechanisms. Below and in the appendix I state and prove the dynamic extension, building on Ely, Frankel and Kamenica (2013). The dynamic obfuscation principle implies that there is no loss of generality in directly choosing a *stochastic process* for the agent's beliefs provided that a dynamic version of the law of total probability (the martingale property) holds *plus* a condition that accounts for the autonomous drift in the agent's beliefs due to probabilistic state transitions.

Obfuscation Principle Indeed, the obfuscation principle allows us to apply dynamic programming to derive the optimal policy. In particular, it implies that there is an optimal value function for the principal that depends only on the agent's belief about the current state.[15] Moreover the optimal value function is characterized by "one-shot" deviations in which, starting from any current belief, the principal chooses a distribution

conditional on state $s = 1$ and $s = 0$ respectively, where $\alpha = (2\mu_t - 1)/(1 - \mu_t)$. Then by Bayes' rule, message m induces belief

$$\frac{\mu_t f_1(m)}{\mu_t f_1(m) + (1-\mu_t) f_0(m)} = \frac{m^{\alpha+1}}{m^{\alpha}} = m.$$

[15] In particular that the optimal continuation value does not depend on the current state of the process s_t. See the discussion following Lemma 2 below.

over subsequent beliefs ν to maximize the expected value of the associated flow payoffs $u(\nu)$ plus (discounted) continuation payoffs.

Feasible deviations must satisfy two constraints. The first is familiar from static problems: the distribution of induced beliefs ν must have expectation equal to the agent's current belief μ_t. The second is new to the dynamic framework: there is a continuation value function that summarizes future payoffs from inducing ν. This function is evaluated at a new belief which incorporates further updating by the agent based on the possibility of state transitions in the intervening time.

Let us illustrate these points in the basic model of this section and show that the optimal policy is a beep with delay t^*. First we will derive its associated value function. As discussed in the introduction, under the mechanism the agent's beliefs are allowed to drift upward to the threshold p^* before beeps begin arriving at rate λ. Let us write $\tau(\mu)$ for the length of time required for the beliefs to drift from $\mu < p^*$ to threshold.[16] Starting from belief μ, the principal will earn a flow payoff of 1 until the beliefs reach p^*, i.e. for duration $\tau(\mu)$. After that, the principal continues to earn a flow payoff of 1 until the first beep sounds. Thus for beliefs $\mu \leq p^*$ the delayed beep earns the principal the following discounted expected continuation value.

$$V(\mu) = \frac{1}{r}\left[(1 - e^{-r\tau(\mu)}) + e^{-r\tau(\mu)}\left(\frac{r}{r+\lambda}\right)\right]. \tag{2}$$

This value function is depicted in Figure 1. In the figure I have extrapolated $V(\cdot)$ linearly for beliefs strictly between p^* and 1. (The continuation value at $\mu = 1$ must be zero under any policy because once the agent is certain that the state is $s = 1$, he will be

[16] The equation $\mu + (1-\mu)(1 - e^{-\lambda\tau(\mu)}) = p^*$ defines $\tau(\mu)$, thus

$$\tau(\mu) = -\frac{1}{\lambda}\log\left(\frac{1-p^*}{1-\mu}\right).$$

certain forever thereafter.) Since the delayed-beep policy never induces a belief in that region this is effectively a "guess" of the optimal value over that range and this guess will be verified in the analysis below.[17]

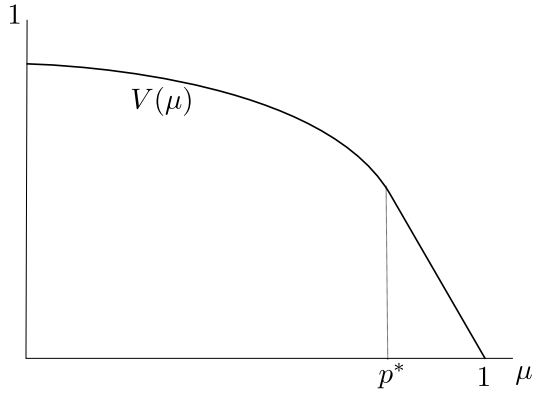


Figure 1: Value function for the t^*-delayed beep. It is strictly concave to the left of p^* and linear to the right.

Next, to show that this is the optimal value, lets consider one-shot deviations from the policy. Fix a small time interval $h > 0$ and consider the payoff to the principal from inducing belief ν and and then reverting to the original mechanism h units of time later. This payoff is the sum of the short-run flow payoff, approximately $u(\nu) \cdot h$, and the discounted continuation value. To calculate the discounted continuation value, note that $(1 - e^{-\lambda h})$ is the probability that an email arrives in the time interval h, and so the agent's belief will be updated from ν to

$$f(\nu) = \nu + (1 - \nu)\left(1 - e^{-\lambda h}\right)$$

which, to a first-order approximation, is $f(\nu) \approx \nu + \frac{d\nu}{dt} \cdot h$ (see Equation 1.) The relevant continuation value for the principal is therefore $V(f(\nu))$, and the overall payoff to the

[17] Sending the agent to such a belief is one possible deviation from the delayed beep policy. In order to test whether such a deviation is an improvement we must know what the continuation value would be after such a deviation.

principal from inducing ν is

$$u(\nu)h + e^{-rh}V(f(\nu)) + O(h^2).$$

The above gives the value from inducing the particular belief ν, but by the obfuscation principle feasible policies induce a *distribution* $q \in \Delta(\Delta S)$ over induced beliefs whose expectation equals the original belief μ. The optimal value $V(\mu)$ should equal the maximum expected payoff attainable from any such q.

$$V(\mu) = \max_{\substack{q \in \Delta(\Delta S) \\ \mathbf{E}q = \mu}} \mathbf{E}_q \left[u(\nu)h + e^{-rh}V(f(\nu)) \right] + O(h^2).$$

Finally, in the limit[18] as h approaches zero, we obtain the Hamilton-Jacobi-Bellman (HJB) equation for the principal's dynamic program:

$$rV(\mu) = \max_{\substack{q \in \Delta(\Delta S) \\ \mathbf{E}q = \mu}} \mathbf{E}_q \left[u(\nu) + V'(\nu)\frac{d\nu}{dt} \right]. \tag{3}$$

Concavification The derived HJB equation states that the optimal policy chooses a distribution over induced beliefs to maximize the expected flow payoff plus the expected rate of change of the continuation value. Following Kamenica and Gentzkow (2011) and Aumann and Maschler (1995), the particular form of the constraint set ($\mathbf{E}q = \mu$) implies that the HJB equation can be given a convenient geometric interpretation:

$$rV = \text{cav}\left[u + V' \cdot \frac{d\nu}{dt} \right]. \tag{4}$$

The value function is the *concavification* (cav) of the flow payoff function plus the time-derivative of the value function itself. The concavification is defined as the pointwise

[18] The online appendix has the formal derivation.

smallest concave function which is no smaller than the function being concavified.[19]

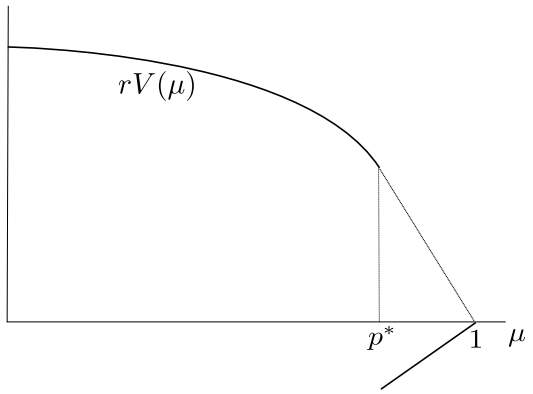


Figure 2: Concavification. The solid line is the function $u + V' \cdot \frac{dv}{dt}$. To the left of p^* it equals rV. When the function is concavified we add the linear segment thus recovering rV globally.

Concavification not only characterizes the optimal value function, it also reveals and explains the optimal policy. To see this, first consider beliefs μ such that the concavified objective coincides with the original function. This implies that the optimum in Equation 3 is a degenerate lottery: it is optimal to reveal no information. On the other hand, when the concavification at μ is strictly higher the optimum is achieved by a randomization over beliefs, i.e. an informative message.

It is this geometric approach that we will use to verify the optimal value function and policy. The bracketed expression in Equation 4 is plotted as the solid curve in Figure 2. At beliefs μ to the left of p^*, it is given by[20] $\left(1 - e^{-r\tau(\mu)}\right) + e^{-r\tau(\mu)} \left(\frac{r}{r+\lambda}\right)$ which is simply $rV(\mu) > 0$. To the right of p^* (where the flow payoff u is zero) the expression in brackets is strictly negative because $V' < 0$. Note that the segment to the left of p^* is already strictly concave so that concavification of the overall function simply adds the dashed linear segment. The resulting function is equal to rV and thus

[19] Intuitively, choosing q in Equation 3 is equivalent to choosing a convex combination of points on the graph of the function. The collection of all such points is the convex hull of the graph and the the maximum in Equation 3 is the upper boundary of that set, a concave function.

[20] See the online appendix for detailed calculations.

we have verified that V, the value function associated with the delayed beep, solves the the HJB Equation 4.

This shows that the delayed beep is optimal and moreover the geometric analysis explains why. Since rV is strictly concave to the left of p^*, it is optimal to reveal no information when the agent's beliefs are in this range. When the agent receives no information his beliefs drift upward and reach p^* at time t^*. Thus, the optimal policy is to reveal no information for the initial period of length t^*. A beep with delay t^* achieves this. Next, at belief p^* (or higher) the concavification is strictly higher than the original function implying that the optimal mechanism sends an informative message. Indeed the concavified value is simply the convex combination of the values at p^* and at $\mu = 1$. That is, the optimal message induces those two beliefs (with the respective probabilities that satisfy the law of total probability.) As argued in the introduction, at date t^* and after, the delayed beep does exactly this.

Discussion This section has introduced the main ideas through a simple example. The analysis was in continuous time. In the next section I generalize the model to any number of actions, any number of states, a general Markov transition process, and general payoffs, including state-dependent payoffs for both principal and agent. I also show how to analyze the model in discrete time. The discrete time analysis has several advantages. First, in the derivation above I "guessed" the optimal value function and verified that it was a fixed point of the HJB equation. In discrete time I show that the value function is a fixed point of a contraction mapping. This means in particular that the solution can be derived algorithmically by value iteration. Moreover uniqueness of the fixed point follows immediately. In addition, each of the steps in the iteration has a geometric interpretation which facilitates the analysis and adds intuition. Finally, the continuous time optimal mechanism can be recovered by taking limits.

3 General Single Agent Model

In the general model with a single agent there is a finite set S of states, a finite set A of actions. The evolution of the state follows a continuous time stochastic process described as follows. State transitions arrive at rate $\lambda > 0$. Conditional on a transition at date t, the new state is drawn from a distribution $M_s \in \Delta S$ where s is the state prior to the transition.

A discrete time formulation will be used to analyze the general model. As will be shown, an advantage of discrete time is that it enables an iterative algorithm to compute the optimal mechanism. Then in specific examples the continuous time optimal mechanism can be found by taking limits as the period length shrinks.

The discrete-time order of events is as follows. The agent enters period t with beliefs μ_t. The principal then observes the current state s_t and sends a message to the agent. In response to the message the agent updates to an *interim* belief ν_t and takes an action. All of the preceding can be regarded as happening instantaneously. Next a discrete length of time passes and we enter period $t+1$.

In the time interval between periods the agent is aware of the possibility of state transitions and further updates his beliefs to take these into account. According to the following lemma (whose proof is in the Online Appendix), the new belief, μ_{t+1} is given by a linear function f of the previous interim belief ν_t.

Lemma 1. *There is a linear map f such that $\mu_{t+1} = f(\nu_t)$.*

The agent's choice of action a_t maximizes the expected value of her state-dependent utility function x, $a_t \in \text{argmax}_a \mathbf{E}_{\nu_t} x(a, s)$.

The principal's payoff $u(a, s)$ depends on both the state s and the agent's action and therefore indirectly depends on the agent's interim belief, i.e. $u(\nu, s) = u(a(\nu), s)$. Henceforth we will take the indirect utility function $u(\nu, s)$ as a primitive and assume

only that it is bounded and upper semi-continuous.[21] The long-run average discounted payoff of the principal is then $(1-\delta)\sum_{t=0}^{\infty}\delta^t u(\nu_t, s_t)$, where δ is the discrete-time discount factor.

A policy for the principal is a rule $\sigma(h_t) \in \Delta M_t$ which maps the principal's complete prior history h_t into a probability distribution over messages. The message space M_t is unrestricted. The principal's history includes all past and current realizations of the process, all previous messages, and all actions taken by the agent.

The set of all policies is unwieldy for purposes of optimization. The obfuscation principle, stated below and proven in the Appendix, allows us to reformulate the problem into an equivalent one in which instead of choosing a policy, the principal is directly specifying a stochastic process for the agent's beliefs.[22] [23]

Lemma 2 (The Obfuscation Principle)**.** *Any policy σ induces a stochastic process (μ_t, ν_t) satisfying*

1. $\mathbf{E}(\nu_t \mid \mu_t) = \mu_t$,
2. $\mu_{t+1} = f(\nu_t)$.

The principal's expected payoff from such a policy is

$$\mathbf{E}\sum_{t=0}^{\infty}\delta^t u(\nu_t) \tag{5}$$

where $u(\nu_t) := \sum_{s\in S}\nu_t(s)u(\nu_t, s)$ and the expectation above is taken with respect to the stochastic process ν_t.

[21] When the agent is maximizing the expected value of x, there will be interim beliefs at which the agent is indifferent among multiple actions. Assuming upper-semicontinuity of u is equivalent to breaking ties in favor of the principal.

[22] The obfuscation principle is conceptually different from the revelation principle. The revelation principle shows that any feasible mechanism can be replaced by a direct revelation mechanism. With the obfuscation principle we don't know in advance that the stochastic process is feasible. We show the feasibility by constructing an appropriate direct obfuscation mechanism.

[23] It would be enough to work just with μ_t or just ν_t below but keeping separate notation for each eases exposition.

Conversely any stochastic process (μ_t, ν_t) with initial belief μ_0 satisfying the above properties can be generated by a policy σ which depends only on the current belief μ_t and the current state s_t, i.e. $\sigma(h_t) = \sigma(\mu_t, s_t)$.

This is a conceptually straightforward extension of analogous results in Aumann and Maschler (1995), Kamenica and Gentzkow (2011), and Ely, Frankel and Kamenica (2013). The proof, which has to contend with potentially infinite message spaces and histories, proceeds somewhat indirectly and is in the Appendix. There is a simplifying benefit of the obfuscation principle that is new to the dynamic setting. Note that according to Equation 5, the current state s_t can be "integrated out" of the principal's objective function. That is, two histories which lead to the same belief for the agent μ_t lead to the same continuation value even if they differ in the realizations of the state s_t. This is true even though the future evolution of the process differs following these histories and the principal could conceivably capitalize on this difference by using different continuation policies. The proof of the obfuscation principle entails showing that there is no benefit from doing so and it is without loss to pool all such histories together as a single state in the dynamic program.

With these preliminaries in hand, we can now solve the principal's optimization problem. By the obfuscation principle, the current μ_t is a state variable for the principal's dynamic optimization. By the principal of optimality, he chooses a lottery over interim beliefs ν_t and earns the flow payoff $u(\nu_t)$. Then the belief is updated to $\mu_{t+1} = f(\nu_t)$ and the principal earns the associated discounted optimal continuation value. The Bellman equation is as follows.

$$V(\mu_t) = \max_{\substack{q \in \Delta(\Delta S) \\ \mathbf{E}q = \mu_t}} \mathbf{E}_q \left[(1-\delta) u(\nu_t) + \delta V(f(\nu_t)) \right]$$

or simply

$$V = \text{cav}\left[(1-\delta)u + \delta\left(V \circ f\right)\right]. \tag{6}$$

where cav denotes concavification. Relative to the static persuasion problem in Kamenica and Gentzkow (2011), the novelty that arises in the dynamic setting is that the value function is the concavification of a function that involves the value function itself. Fortunately this fixed-point problem can be solved in a conceptually straightforward way when we make two observations. First, the right-hand side can be viewed as a functional operator mapping a candidate value function into a re-optimized value function. By standard arguments this operator is a contraction and therefore has a unique fixed point which can be found by iteration. The proof is in the Online Appendix.

Theorem 1. *The optimal value function is characterized by the functional equation in Equation 6. In particular, V is the unique fixed point of the operator*

$$TV = \text{cav}\left[(1-\delta)u + \delta\left(V \circ f\right)\right].$$

which is a contraction mapping and therefore converges by iteration to V.

Second, the set of operations on the right-hand side all have convenient geometric interpretations (composition, convex combination, concavification) making this iteration easy to visualize and interpret. As an illustration, in the online appendix we solve several examples with just a series of diagrams.

4 Example: Selling A Depreciating Good.

Consider a monopolist selling a product of depreciating quality. The product is initially of high quality but at a random time observed only by the monopolist the product de-

preciates and the quality is thereafter low.[24] The time of the transition is exponentially distributed with parameter $\lambda = 1$. The monopolist's information policy consists of public disclosures at each time t conveying information about the current quality. At each time t there is a flow demand $Q(p_t, \nu)$ as a function of the price p_t set by the monopolist, and the current belief ν about the quality of the product, i.e. the probability that the quality remains high. Absent any information from the monopoly, this belief is trending downward, i.e. $\frac{d\nu}{dt} = -\lambda\nu$. Production costs are zero. The monopolist chooses an information policy as well as a pricing strategy to maximize expected discounted total profits, with discount rate $r = 1$.[25] The detailed calculations for this section are in the Online Appendix.

Quality-Senstive Demand Consider demand given by $Q(p_t, \nu) = \nu - p_t$. When the monopolist's information policy has induced belief ν, the profit maximizing price (setting marginal revenue to zero) is $p_t = \nu/2$ yielding profit $u(\nu) = \nu^2/4$. This profit function is convex in the induced belief ν. In the case of a strictly convex $u(\nu)$, it is easy to show that because the concavification of any convex function is linear, the optimal value function is linear and therefore that the optimal mechanism is full information.

Quality-Insensitive Demand Consider demand given by $Q(p_t, \nu) = \nu^{1/4} - p_t$. Demand is less sensitive to quality than in the previous example. The profit-maximizing price is now $\nu^{1/4}/2$ yielding profit function $u(\nu) = \nu^{1/2}/4$. This profit function is strictly concave. It is easy to show that a strictly concave profit function generates a strictly concave value function and therefore an optimal policy of no information.

[24] The example admits a few different interpretations. We could think of a periodic rental market for a depreciating durable good or just a production technology that is subject to depreciation. The demand at a given time t could represent a single (short-lived) consumer with a private type or a large market of heterogenous consumers.

[25] Since the optimal price will be a one-to-one function of the belief, one implementation of the policy has the price itself act as the signal.

Mixed Demand Finally consider a market consisting of both quality-sensitive and quality-insensitive consumers. The demand $Q(p_t, \nu) = \nu + \nu^{1/4} - 2p_t$ aggregates the demand schedules from the two previous examples. It yields monopoly price $(\nu + \nu^{1/4})/4$ and profit $u(\nu) = (\nu + \nu^{1/4})^2/8$. This profit function is concave and then convex as ν ranges from 0 to 1. The optimal information policy is partially informative. The monopolist makes an announcement when the product depreciates but also makes this announcement with positive probability when it has not. The updated belief after the announcement, call it ν^*, is thus strictly between 0 and 1. Following the announcement the optimal policy is silent and allows the belief to drift toward 0.[26] The optimal value function is depicted in Figure 3. The function W gives the value of the non-disclosure policy, it has the same shape as u. The value function coincides with W below ν^*. Above ν^* it is linear reflecting that it is optimal to reveal a partially informative message inducing either ν^* or $\mu = 1$.

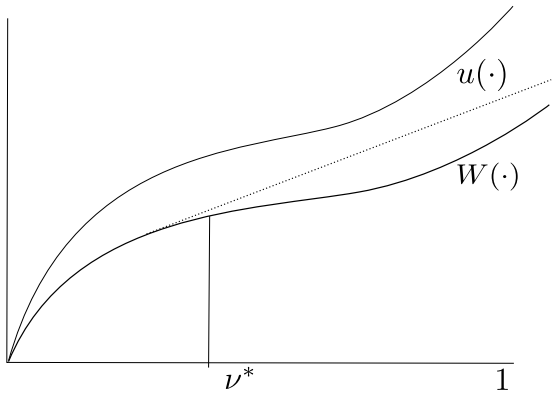


Figure 3: Value function for the mixed-demand case.

[26] Renault, Solan and Vieille (2014) discuss "greedy" policies: those which simply apply the static-optimal policy at every state. Often the greedy policy is in fact optimal in the dynamic problem. Indeed this will be the case whenever the flow payoff function u is globally concave or globally convex as in the case of quality-sensitive and quality-insensitive demands. On the other hand, in the mixed-demand case the optimal policy is not greedy. Indeed the dynamic optimal policy is less informative than the static optimum. This occurs because a patient monopolist has an overall flatter value function due to the fact that future low payoffs (after depreciation) are internalized at early states. The cost of a less informative announcement is that such an announcement must arrive faster, but with a flatter value function this cost is reduced.

5 Multiple Agents

Often a planner who controls information can use it not just to influence the beliefs of individuals, but also to facilitate or disrupt the coordination of a set of individuals. In this section we consider a simple example of a dynamic multi-agent persuasion problem. The framework is a stylized model of a bank run. The bank (the principal) strategically releases information about its solvency in order to incentivize depositors (the agents) to keep their deposits at the bank. The new ingredient in the multi-agent setting is that depositors' incentives to withdraw depend not just on the underlying exogenous state (the health of the bank) but also the behavior of the other agents.

A number of qualitative conclusions emerge from this extension. First, there is a trade-off between public and private signals. This trade-off in general depends on the nature of the strategic interaction among agents and the objective of the principal.[27] In the bank run context the principal wants to impede coordinated action by the agents and this is achieved using private signals generating asymmetric information.

Second, maximal mis-coordination is achieved by inducing negative correlation in the agents' beliefs. However, Bayesian updating together with incentive constraints imply a bound on how negatively correlated two agent's beliefs can be. The problem thus reduces to characterizing this bound and finding a mechanism that attains it.

Finally, maximal mis-coordination can be achieved without compromising on the incentives of individual agents. An *individually optimal* mechanism is one which maximizes for each agent separately the expected time before withdrawal. In the bank run example considered here there exists an overall optimal mechanism that is also individually optimal. In particular, the optimal mechanism is found by first characterizing

[27] As is well-known from static models (for example Morris and Shin (2002)) coordinated action by the agents is facilitated by common belief in the underlying state of the world. We can thus think of the principal's problem as controlling the flow of information to minimize common belief, i.e. interfere with *common learning*, see Cripps et al. (2008) and Steiner and Stewart (2011).

individually belief paths and then finding the joint process in which these belief paths are maximally mis-coordinated.

The setup for the example is the following. There are two agents and each has a deposit with the bank. The bank may be either healthy or distressed. Initially the bank is healthy and it transitions to the distressed state at Poisson rate λ. The current state is observed only by the bank. Each agent wishes to withdraw his deposit, i.e. *run*, if the bank is distressed. However, even if the bank is healthy each agent nevertheless wishes to run if he expects that the other agent will run. Withdrawal is an irreversible decision, observable by the bank but unobservable to the other depositor. We consider the following parameterization[28][29] of the agent's payoffs: The objective of the bank is

	wait	run
wait	p^*	p^*
run	0	1

healthy

	wait	run
wait	p^*	p^*
run	1	1

distressed

to forestall a bank run, defined here as the date at which both agents have withdrawn. The bank may also suffer a loss during periods in which only a single depositor has withdrawn. However as we will show below it will be unnecessary to formally model these losses. The mechanism which maximally delays a bank run does so by also maximally delaying the date at which any single depositor withdraws.

[28]Under this specification the strategic interaction of the agents is particularly simple: each depositor wishes to withdraw if and only if the total probability exceeds p^* that *either* the bank is distressed or the other agent has already withdrawn. This facilitates a direct comparison with the single-agent problem in which p^* was the key belief threshold for the agent.

[29] Adding a constant to any column does not change the incentives of a depositor and therefore equilibrium behavior. Therefore the fact that, for example the players earn higher payoffs from (`run`, `run`) than (`wait`, `wait`) is just a normalization. On the other hand, that the payoff gain from `run` rather than `wait` is state-independent is a real simplification. If these were different the optimal information policy would involve two parameters rather than the single parameter p^*.

Mechanisms The bank commits to an information policy. Just as in the single-agent case there is no restriction on the set of feasible policies: the messages sent to each agent at each point in time can depend in arbitrary ways on the past history. In addition the messages sent to depositor 1 can be designed to provide arbitrary information about the current or past messages sent to depositor 2 and vice versa. For example, depositor 1 can be told what depositor 2 believes about the health of the bank, whether or not depositor 2 has withdrawn, what depositor 2 has been told about depositor 1, etc. And all of these pieces of information can be randomized, delayed, or otherwise related to the underlying process.

A simple benchmark is a *public* mechanism in which all signals are observed by both agents. In any public mechanism the beliefs of the depositors must be the same at each point in time and in particular their withdrawal times are perfectly correlated. Thus, if restricted to a public mechanism the bank's optimization problem would be equivalent to maximizing the delay before a single agent's belief exceeds p^*. It follows from our previous results that the optimal public mechanism is a delayed public signal and the expected time before the onset of the bank run is $t^* + 1/\lambda$.

The bank can do better with private messages. Private messages can induce asymmetric beliefs and thereby reduce the correlation in withdrawal times. For any given *marginal* distributions of withdrawal times for each agent individually, the less correlated is the *joint* distribution the better off is the principal.

However, once we move beyond the class of public mechanisms, the obfuscation principle no longer suffices to characterize all feasible mechanisms and there is no useful generalization for the multi-agent case.[30] Therefore, the analysis that follows will instead employ a version of the revelation principle. It is without loss of generality to

[30]For example, suppose that the depositors initially share common knowledge that the bank is distressed with probability q. Any implementable *joint* distribution over posteriors must have expectation q for each depositor separately. However, not all such joint distributions are implementable.

consider direct mechanisms of the following form. The principal, possibly randomly, selects withdrawal times (r_1, r_2) and when date r_i is reached, privately instructs depositor i to withdraw. Indeed any feasible mechanism is characterized by a joint distribution of times $(\tilde{r}_1, \tilde{r}_2)$ such that it is incentive compatible for depositor i to withdraw when and only when instructed.

The first proposition gives an upper bound on the expected delay before a bank run based on a relaxed problem in which strategic incentive constraints are ignored.[31]

Proposition 1. *Define*

$$t^{**} = t^* + \frac{1}{\lambda}\log(1 + p^*).$$

Under any incentive-compatible mechanism, a bank-run cannot be delayed longer than t^{**}. Indeed, the probability that a bank run occurs on or before any given date t is at least

$$G(t) = \begin{cases} 0 & \text{if } t \leq t^{**} \\ 1 - e^{-\lambda(t - t^{**})} & \text{if } t > t^{**}. \end{cases} \tag{7}$$

There exists an incentive compatible direct mechanism that achieves this bound and is also *individually optimal*: each agent waits on average the maximal delay time $t^* + 1/\lambda$. In particular the bank achieves the maximum expected delay before a bank run while simultaneously maximally delaying the expected withdrawal time of each agent.[32]

Theorem 2. *There exists an incentive compatible direct mechanism $(\tilde{r}_1, \tilde{r}_2)$ such that for each date $t > t^*$,*

[31] I thank Stephen Morris for a discussion that led to this approach.

[32] Since the date of the bank run is a property of the *joint* distribution (i.e. $\max\{\tilde{r}_1, \tilde{r}_2\}$) and all individually optimal mechanisms achieve the same (optimal) *marginal* distributions for $\tilde{r}_1$ and $\tilde{r}_2$ this is related to the well-studied mathematical problem of characterizing the range of joint distributions consistent with given marginals. See Nelsen (1999).

1. *The probability of a bank run prior to t is equal to the lower bound $G(t)$.*

2. *The probability that depositor i withdraws prior to t is $1 - e^{-\lambda(t-t^*)}$.*

In particular the distributions of $\tilde{r}_i$ and $\max \tilde{r}_i$ first-order stochastically dominate all incentive compatible mechanisms.

Note that the first-order dominance implies that this mechanism is optimal for any payoff function of the bank. If l_1 is the flow loss when a single depositor has withdrawn and l_2 is the flow loss when both agents have withdrawn (a bank run), then the mechanism is optimal as long as $l_2 \geq l_1$.

6 Conclusion

This paper develops methods for studying dynamic information management in single and multi-agent scenarios. The framework has a number of special features which are worth discussing further. The behavior of the agent has mostly been modeled as a black-box response to current beliefs. This can be interpreted literally as a myopic agent maximizing short-run payoffs or alternatively as a sequence of short-run agents. However it should be emphasized that in many settings the agent's action is not observable or contractible. When the mechanism cannot condition on the agent's action it is optimal for the agent to maximize payoffs period-by-period even if the agent has long-run incentives.

Going further, if the agent has long-run incentives and his actions are observable there is in principle the possibility of using future information revelation as the basis for promises and threats to further incentivize the agent. However these incentives are very limited in scope. Indeed a working paper version (Ely (2015)) considers an extension to the basic beeps example in which the agent is patient. In this setting dynamic

incentives are of no additional value to the principal and the optimal mechanism is identical to the one studied here. A general study of dynamic persuasion with long-run incentives is a topic of ongoing research.

In this paper the stochastic process for the underlying state is assumed to follow Poisson transitions. This allows simple closed-form expressions but is not essential for the analysis. The role of the underlying process in the optimal mechanism is completely summarized by the law of motion f, and any process which gives rise to a time-invariant (and well-behaved) law of motion can be accommodated.

The agent is assumed to have no access to external information about the state and to rely only on communication from the principal. Exogenous *public* information can be accommodated in a straightforward way. This is because the principal can compute the agent's updated belief after each realization of a public signal, and adapt his message to the new belief. Indeed the information acquired can even be allowed to depend on the action choice of the agent, allowing for example application to problems of public experimentation. In particular we can continue to use this updated belief of the agent's as a state variable for dynamic programming because it will completely summarize the expected continuation payoff of the principal.

However this is no longer true if the agent has access to private information about the state. Then the agent's beliefs, and therefore his actions, can depend on all past realizations of this private information. A (direct) mechanism would now involve the agent reporting (in an incentive compatible way) his private information to the principal and the principal responding with a message of his own. In a static setting Kolotilin et al. (2015) study persuasion of a privately-informed receiver and show that there is no loss to the principal from a policy that does not condition on the agent's reports. It is an open question whether this insight can also simplify the analysis of dynamic problems with private information.

The principal's commitment power is important here as it is in static information design settings. Formally, it enables the obfuscation principle to pool all histories leading to the same beliefs and "integrate out" the underlying state. It is worth emphasizing that even in contexts in which the principal is unable to fully commit, the commitment solution is useful as a benchmark representing a "first-best" policy that second-best policies can be compared to. Finally, in the absence of commitment, the dynamic setting raises the new possibility that repeated-game arguments can enable the principal to approach or even achieve the value of commitment. This is an interesting direction for future research.

References

Angeletos, George-Marios, Christian Hellwig, and Alessandro Pavan. 2007. "Dynamic global games of regime change: Learning, multiplicity, and the timing of attacks." *Econometrica*, 75(3): 711–756.

Aoyagi, Masaki. 2010. "Information feedback in a dynamic tournament." *Games and Economic Behavior*, 70(2): 242–260.

Aumann, Robert J, and Michael Maschler. 1995. *Repeated games with incomplete information.* MIT press.

Bergemann, Dirk, and Stephen Morris. 2013. "Robust predictions in games with incomplete information." *Econometrica*, 81(4): 1251–1308.

Bergemann, Dirk, and Stephen Morris. 2014. "The comparison of information structures in games: Bayes correlated equilibrium and individual sufficiency."

Bergemann, Dirk, Benjamin Brooks, and Stephen Morris. 2014. "Extremal Information Structures in First-Price Auctions." Princeton University.

Billingsley, Patrick. 2008. *Probability and measure.* John Wiley & Sons.

Che, Yeon-Koo, and Johannes Hörner. 2013. "Optimal Design for Social Learning." Mimeo.

Cripps, Martin W, Jeffrey C Ely, George J Mailath, and Larry Samuelson. 2008. "Common learning." *Econometrica*, 76(4): 909–933.

Ely, Jeffrey, Alexander Frankel, and Emir Kamenica. 2013. "Suspense and surprise." mimeo, University of Chicago.

Ely, Jeffrey C. 2015. "Beeps (working paper)."

Gershkov, Alex, and Balázs Szentes. 2009. "Optimal voting schemes with costly information acquisition." *Journal of Economic Theory*, 144(1): 36–68.

Halac, Marina, Navin Kartik, and Qingmin Liu. 2014. "Contests for Experimentation."

Kamenica, Emir, and Matthew Gentzkow. 2011. "Bayesian persuasion." *American Economic Review*, 101(6).

Kolotilin, Anton, Ming Li, Tymofiy Mylovanov, and Andriy Zapechelnyuk. 2015. "Persuasion of a Privately Informed Receiver." Working paper.

Kremer, Ilan, Yishay Mansour, and Motty Perry. 2013. "Implementing the" Wisdom of the Crowd"."

Lizzeri, Alessandro, Margaret A Meyer, and Nicola Persico. 2002. "The incentive effects of interim performance evaluations."

Morris, Stephen, and Hyun Song Shin. 1999. "A theory of the onset of currency attacks." *The Asian Financial Crisis: Causes, Contagion and Consequences*, 2: 230.

Morris, Stephen, and Hyun Song Shin. 2002. "Social value of public information." *The American Economic Review*, 92(5): 1521–1534.

Nelsen, Roger B. 1999. *An introduction to copulas.* Springer.

Orlov, Dmitry. 2013. "Optimal Design of Internal Disclosure."

Renault, Jérôme, Eilon Solan, and Nicolas Vieille. 2014. "Optimal Dynamic Information Provision." *arXiv preprint arXiv:1407.5649*.

Steiner, Jakub, and Colin Stewart. 2011. "Communication, timing, and common learning." *Journal of Economic Theory*, 146(1): 230–247.

Taneva, Ina. 2014. "Information design." *Unpublished manuscript*.

A Appendix

Proof of the Obfuscation Principle. That a policy induces a stochastic process satisfying the two conditions is a direct implication of Bayesian updating by the agent based on knowledge of the policy used by the principal. The first condition is simply the law of total probability and the second condition reflects the agent's knowledge of the underlying stochastic process governing the state.

Now consider any (possibly indirect) mechanism. Let h denote an infinite history consisting of realized states, messages, and actions in each period. Write $\nu_t(h)$, $s_t(h)$, and $a_t(h)$ for the date t interim beliefs, realized states and actions along history h. The principal's payoff from the mechanism is $\mathbf{E}_h \sum_{t=0}^{\infty} \delta^t u(a_t(h), s_t(h))$ where the expectation is taken with respect to the distribution of infinite histories induced by the mechanism. If we write $H_t(s) = \{s_t(h) = s\}$ for the set of all histories for which s is the state at date t, then recalling that the agent's optimal action is a function of ν, $a(\nu)$, this

payoff can be re-written $\mathbf{E}_\nu \sum_{t=0}^{\infty} \delta^t \sum_{s \in S} \text{Prob}(H_t(s) \mid \nu_t) u(a(\nu_t), s)$ where the expectation is taken with respect to the stochastic process for interim beliefs ν_t associated with the mechanism. Since ν_t is the Bayesian posterior given the policy and all information previously revealed, $\text{Prob}(H(s) \mid \nu_t) = \nu_t(s)$ and the payoff reduces to $\mathbf{E}_\nu \sum_{t=0}^{\infty} \delta^t u(\nu_t)$.

To prove the converse, let (μ_t, ν_t) be any stochastic process satisfying item 1, and item 2. We will construct a policy which generates it and which depends only on the current belief μ_t and the current state s_t. Fix t and let Z denote the conditional distribution of ν_t given μ_t. The policy is a *direct obfuscation mechanism* in which the principal tells the agent directly what his beliefs should be. To that end, let the message space be $M_t = \Delta(S)$. Let $\sigma_s \in \Delta(M)$ denote the lottery over messages when the current belief is μ_t and the current state is $s_t = s$. The probability σ_s is defined by the following law: for measurable $B \subset \Delta(M)$,

$$\sigma_s(B) = \int_{\nu \in M} \frac{\nu(s)}{\mu_t(s)} dZ(\nu). \tag{8}$$

That is, σ_s is defined to be absolutely continuous with respect to Z with Radon-Nikodym derivative equal to $\frac{\nu(s)}{\mu_t(s)} : \Delta(S) \to \mathbf{R}$. So defined, σ_s is a probability because it is non-negative, countably additive and for any measurable $B \subset \Delta(M)$, we have $\int_{\nu \in B} \nu(s) dZ \leq \int_{\nu \in \Delta(M)} \nu(s) dZ = \mathbf{E}\nu(s)$ and the latter is equal to $\mu_t(s)$ by item 1. Thus, the right-hand side of Equation 8 is less than or equal to 1 and equal to 1 when $B = \Delta(M)$.

From the point of view of the agent, who does not know the current state s but knows that the policy is σ_s and has beliefs μ_t about s, the total probability of a set $B \in M$ is

$$\sum_s \mu_t(s) \sigma_s(B) = \sum_s \mu_t(s) \int_{\nu \in B} \frac{\nu(s)}{\mu_t(s)} dZ(\nu) = \sum_s \int_{\nu \in B} \nu(s) dZ(\nu) = \int_{\nu \in B} 1 dZ(\nu) = Z(B)$$

Thus, the policy generates the desired conditional distribution over messages. It remains to show that when the principal uses the policy and the agent observes message

ν his posterior beliefs about s_t are indeed equal to ν. Fix a state s, consider the probability space $(\Delta(S), Z)$ and defined over it the random variable given by $y(\nu) = \nu(s)$. By construction, for all $B \in \Delta(S)$, we have $\int_{\nu \in B} y(\nu) dZ(\nu) = \mu_t(s)\sigma_s(B) = \text{Prob}(\{s\} \times B)$ so that y is a version of the conditional probability of s (Billingsley, 2008, Section 33). Thus, $y(\nu) = \nu(s)$ is the Bayesian posterior probability of state s upon receiving the message ν and therefore the agent's interim belief. □

Proof of Proposition 1. We will consider a relaxed problem in which the agents are non-interactive (but the principal's objective is still to maximize $\mathbf{E}(\max \tilde{r}_i)$). In this relaxed problem higher-order beliefs are irrelevant, each agent runs if and only if he assigns probability greater than p^* to `distressed`.

In the relaxed problem there is no loss of generality in restricting attention to individually optimal mechanisms. We will now derive bounds on the distribution of run times $\tilde{r}_i$ for an agent facing an individually optimal mechanism. Since the agent waits only if the probability of `distressed` is below p^*, we have for every t,

$$\text{Prob}(\phi \leq t \mid \tilde{r}_i > t) \leq p^*. \tag{9}$$

We can express the conditional probability on the left-hand side as follows.[33]

$$\begin{aligned} \text{Prob}(\phi \leq t \mid \tilde{r}_i > t) &= \frac{\text{Prob}(\phi \leq t) - \text{Prob}(\tilde{r}_i \leq t)}{\text{Prob}(\tilde{r}_i > t)} \\ &= \left(1 + \frac{\text{Prob}(\phi > t)}{\text{Prob}(\phi \leq t) - \text{Prob}(\tilde{r}_i \leq t)}\right)^{-1} \end{aligned}$$

[33] To understand the first line, observe that the numerator equals $\text{Prob}(\{\phi \leq t\} \cap \{\tilde{r}_i > t)\}$ because $\{\phi \leq t\} \cap \{\tilde{r}_i > t)\} = \{\phi \leq t\} \setminus [\{\phi \leq t\} \cap \{\tilde{r}_i \leq t\}]$ and in an individually optimal mechanism $\{\tilde{r}_i \leq t\} \subset \{\phi \leq t\}$ with probability 1.

and thus we can use the inequality in Equation 9 to obtain

$$\text{Prob}(\tilde{r}_i \leq t) \geq \left(\frac{1}{1-p^*}\right)[\text{Prob}(\phi \leq t) - p^*].$$

The right-hand side is non-negative as soon as $t \geq t^*$. In a solution to the relaxed problem the constraint will bind and hence

$$\text{Prob}(\tilde{r}_i \leq t) = \begin{cases} 0 & \text{if } t \leq t^* \\ \left(\frac{1}{1-p^*}\right)[\text{Prob}(\phi \leq t) - p^*] & \text{otherwise.} \end{cases} \tag{10}$$

Next consider $\text{Prob}(\tilde{r}_2 \leq t \mid \tilde{r}_1 > t)$, i.e. the probability agent 1 assigns to agent 2 having already withdrawn conditional on agent 1 yet to hear a beep. It is given by

$$\frac{\text{Prob}(\tilde{r}_2 \leq t) - \text{Prob}(\max \tilde{r}_i \leq t)}{\text{Prob}(\tilde{r}_1 > t)}$$

because the numerator equals the probability that $\tilde{r}_2 \leq t$ and $\tilde{r}_1 > t$. Recall that in an individually optimal mechanism, when an agent runs he assigns probability 1 to `distressed`. Thus[34] $\text{Prob}(\tilde{r}_2 \leq t \mid \tilde{r}_1 > t) \leq \text{Prob}(\phi \leq t \mid \tilde{r}_1 > t)$, and since by Equation 9, the right-hand side is less than or equal to p^*, we have

$$\frac{\text{Prob}(\tilde{r}_2 \leq t) - \text{Prob}(\max \tilde{r}_i \leq t)}{\text{Prob}(\tilde{r}_1 > t)} \leq p^*$$

implying $\text{Prob}(\max \tilde{r}_i \leq t) \geq \text{Prob}(\tilde{r}_2 \leq t) - p^* \text{Prob}(\tilde{r}_1 > t)$. In an individually optimal mechanism the runtimes $\tilde{r}_1$ and $\tilde{r}_2$ have the same marginal distribution (Equation 10) and thus $\text{Prob}(\max \tilde{r}_i \leq t) \geq (1+p^*)\text{Prob}(\tilde{r}_2 \leq t) - p^*$. Using Equation 10 again we

[34] Since the left-hand side equals $\text{Prob}(\tilde{r}_2 \leq t \cap \phi \leq t \mid \tilde{r}_1 > t) + \text{Prob}(\tilde{r}_2 \leq t \cap \phi > t \mid \tilde{r}_1 > t)$ and the second term is zero in an individually optimal mechanism.

obtain

$$\text{Prob}(\max \tilde{r}_i \leq t) \geq \left(\frac{1+p^*}{1-p^*}\right) \text{Prob}(\phi \leq t) - \frac{2p^*}{1-p^*}.$$

Recall that $t^{**} = t^* + \frac{1}{\lambda}\log(1+p^*)$, allowing the previous inequality to be rewritten $\text{Prob}(\max \tilde{r}_i \leq t) \geq e^{\lambda t^{**}}\text{Prob}(\phi \leq t) + (1 - e^{\lambda t^{**}})$ and since ϕ is distributed exponentially with parameter λ, $\text{Prob}(\max \tilde{r}_i \leq t) \geq e^{\lambda t^{**}}\left(1 - e^{-\lambda t}\right) + (1 - e^{\lambda t^{**}}) = 1 - e^{-\lambda(t-t^{**})}$. □

Proof of Theorem 2. Let ϕ denote the random time at which the bank transitions from `healthy` to `distressed`, and let F denote its cumulative distribution function. The optimal mechanism can be described as a deterministic functions of ϕ. In particular,

$$\tilde{r}_1(\phi) = \phi + t^*$$

$$\tilde{r}_2(\phi) = \begin{cases} \phi & \text{if } \phi \geq t^* \\ \psi(\phi) & \text{if } \phi < t^* \end{cases}$$

where $\psi(\phi) = -\frac{1}{\lambda}\log\left(\frac{F(\phi)}{p^*}\right) + t^*$.

This mechanism is incentive compatible. First note that $\tilde{r}_i \geq \phi$ so that each agent withdraws only if the bank is distressed. Thus each agent strictly prefers to wait whenever the conditional probability that the bank is distressed is no greater than p^*. Note that $\tilde{r}_1 = \phi + t^*$ is just the delayed beep and thus the mechanism is incentive compatible for depositor 1. The mechanism faced by depositor 2, $\tilde{r}_2$ generates the same stochastic process for beliefs and is therefore also incentive compatible. To see this, consider any $t \geq t^*$ and the probability that the bank is distressed conditional on depositor 2 being instructed to wait. This can happen only if $\phi < t^*$ and $\psi(\phi) < t$. The total probability

is $F(\psi^{-1}(t))$. Now

$$t = \psi\left(\psi^{-1}(t)\right) = -\frac{1}{\lambda}\log\left(\frac{F(\psi^{-1}(t))}{p^*}\right) + t^*$$

hence $F(\psi^{-1}(t)) = p^* e^{-\lambda(t-t^*)}$. The conditional probability is therefore

$$\frac{p^* e^{-\lambda(t-t^*)}}{p^* e^{-\lambda(t-t^*)} + (1 - F(t))} = \frac{p^* e^{-\lambda(t-t^*)}}{p^* e^{-\lambda(t-t^*)} + (1-p^*)e^{-\lambda(t-t^*)}}$$

i.e. p^*.

The mechanism achieves the bound G. From Proposition 1 the probability of a bank run before t^{**} is zero. Now consider any $t > t^{**}$ and the event of a bank run within $[t^{**}, t]$. This is the event that $\tilde{r}_1(\phi) < t$, and $\tilde{r}_2(\phi) < t$, i.e. $\phi + t^* < t$ and $\psi(\phi) < t$.[35] The probability of this event is $F(t-t^*) - F(\psi^{-1}(t)) = 1 - e^{-\lambda(t-t^*)} - p^* e^{-\lambda(t-t^*)} = 1 - (1+p^*)e^{-\lambda(t-t^*)} = 1 - e^{-\lambda(t-t^{**})}$ since by Equation 7, $1 + p^* = e^{\lambda(t^{**}-t^*)}$.

The probability that depositor 1 withdraws prior to $t < t^*$ is $F(t-t^*) = 1 - e^{-\lambda(t-t^*)}$. For depositor 2 the probability is $F(t) - F(\psi^{-1}(t)) = 1 - e^{-\lambda t} - p^* e^{-\lambda(t-t^*)} = 1 - e^{-\lambda t^*} e^{-\lambda(t-t^*)} - p^* e^{-\lambda(t-t^*)} = 1 - (e^{-\lambda t^*} + p^*)e^{-\lambda(t-t^*)} = 1 - e^{-\lambda(t-t^*)}$ since by definition $p^* = 1 - e^{-\lambda t^*}$.

Finally since for every t, these probabilities achieve the bounds in Equation 7 and Equation 10, the distributions first-order stochastically dominate all incentive compatible mechanisms. □

[35] We can ignore the first case in the definition of $\tilde{r}_2$ since if $\phi > t^*$ then the condition $\phi < t$ is already implied by $\phi + t^* < t$.