Beeps*

Jeffrey C. Ely †

First Version: December 2013
This version: February 12, 2015

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 apply it to some examples. I also consider extensions to multiple agents where higher-order beliefs matter, to patient agents where dynamic incentives matter, and to continuous-time problems. *Keywords: beeps, 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. I received excellent questions, comments, and suggestions from Gabriel Carroll, Johannes Horner, Moritz Meyer-ter-Vehn, Simon Board, Sergiu Hart, Andrea Prat, Tanjim Hossain, Eilon Solan, Nicolas Vielle, Debraj Ray, Daria Khromenkova and Stephen Morris. This paper began as a March 2012 blog post at cheaptalk.org. The version of the paper you are reading may be obsolete. The most recent version can be found at `jeffely.com/beeps`

†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.

A number of important new applications can be understood using this framework. For example, we may consider the interaction between a CEO who is overseeing the day-to-day operations of a firm and the board of directors which obtains information only through periodic reports from the CEO. Absent any recent reporting the board will become pessimistic and order an audit. Audits are costly and so the CEO must choose the optimal timing of reports in order to manage the frequency of audits.

A 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. 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. If performance evaluations convince the worker that she has made some progress but she is not quite ready for promotion, then in the absence of such a report she will infer either that she is nearly there and can coast to the finish line, or that successes have been sufficiently rare that promotion is out of reach. In either case she works less hard. If the new technology looks promising to the principal or seems to be the unanimous choice of isolated agents, a policy of making this information public entails the downside that silence will make the next agent too pessimistic to engage in socially beneficial exper-

[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).

imentation.

I develop a general model to analyze the costs and benefits of 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 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. 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. Messages that persuade the agent to take desired actions today also alter the path of beliefs in the future. There is thus a tradeoff between current persuasion and the ability to persuade in the future.

To illustrate these ideas consider the following example which will be used throughout the paper. A researcher is working productively at his desk. Nearby there is a computer and distractions in the form of email are arriving stochastically over time. When an email arrives his computer emits a beep which overwhelms his resistance and he suspends productive work to read email and subsequently waste time surfing the web, reading social media, even *sending* email. Fully aware of his vulnerability to distraction how can he avoid this problem and remain productive for longer?

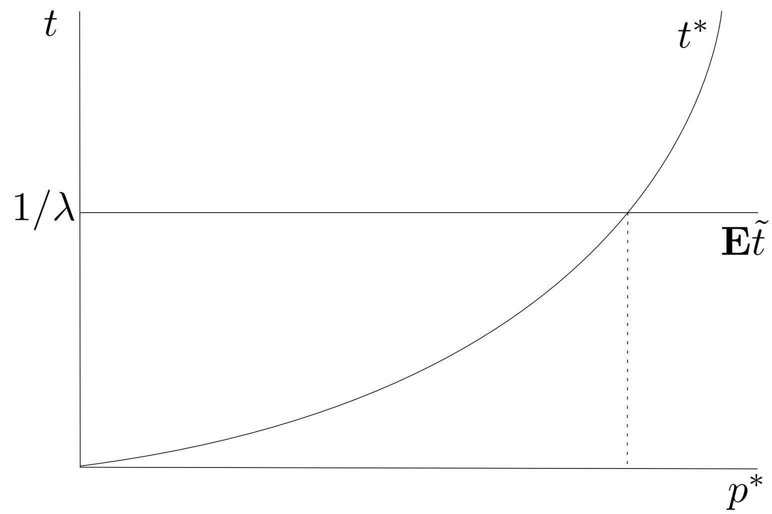
One possibility is to disable the beep. However, there is no free lunch: if he knows that the beep is turned off then as time passes he will become increasingly certain that an email is waiting and he will give in to temptation and check.[3] To formalize this let's suppose that the researcher's degree of temptation is represented by a threshold belief p^* such that once such a time is reached that his posterior belief exceeds p^*, he will stop working. Then, assuming email arrives at Poisson rate λ, turning the beep off affords him an interval of productivity of a certain length

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

as the latter is the time it takes for the first-arrival probability $1 - e^{-\lambda t}$ to reach p^*.

[3] Even if he checks and sees that in fact there is no email there he will still get distracted by the other applications on his computer and lose just as much productivity.

By contrast, when the beep is on, the time spent working without distractions is random and given by the arrival time $\tilde{t}$ of the first beep which causes his posterior to jump discontinuously past p^* to 1. The expected time[4] before the first beep is given by $\mathbf{E}\tilde{t} = 1/\lambda$. The comparison between these two signal technologies is represented in the figure below where the vertical axis measures expected time working as a function of the threshold on the horizontal axis.[5]



Interestingly, a researcher who is easily distracted (represented by a low p^*) should nevertheless amplify distractions by turning the beep on. This is because in return for the distraction, unlike when the beep is off he is able to remain at his desk arbitrarily long when his email beep happens to be silent. The end result is more time on average spent being productive. On the other hand, a researcher who is not so easily tempted is better off silencing her email and benefiting from the relatively long time it takes

[4]The calculations below focus on expected waiting times and thus ignore discounting and possibly non-constant marginal productivity of work over time. Discounting is explicitly included in the general analysis to come. Other sources of non-linearity are interesting extensions to be explored in further work.

[5]Note that the arrival rate of email plays no role in the comparison. This can be understood by considering an innocuous change of time units. A sixty-fold increase in λ is equivalent to rescaling time so that it is measured in minutes rather than seconds. But clearly this won't change the real waiting times nor any comparison between signaling technologies. On the other hand as we will show below the precise details of the *optimal* mechanism will be tailored to λ.

before she can't resist.[6]

Once we observe that filtering information can control the behavior even of an agent who knows he is being manipulated and rationally updates beliefs over time we are naturally led to consider alternative signaling technologies and ultimately the optimal mechanism.

Consider a random beep. In particular suppose that when an email arrives the email software performs a randomization and with probability $z \in (0,1)$ emits a beep. Similar to beep-on (which is equivalent to $z = 1$) she will be induced to check as soon as the first beep sounds. And similar to beep-off ($z = 0$) after a sufficiently long time without a beep she will succumb to temptation and check. It would seem that an interior z combines the worst of both mechanisms but as we have seen any negative incentive effect is coupled with a potentially compensating positive effect. Indeed, the expected waiting time can be calculated as follows.

$$\frac{1 - (1 - p^*)^{-\frac{z}{z-1}}}{\lambda z}$$

For most values of p^*, this expression is non-monotonic in z and hence an interior z is preferred.[7]

Is a random beep optimal among all policies? The random beeps considered above are special because they have false negatives but no false positives. In general it may be optimal to fine tune the randomizations to yield differential false positives and false negatives to achieve a broader set of possible beliefs. Beeps with continuously variable volumes chosen judiciously as a function of history can calibrate beliefs even finer. And in a dynamic framework there are many more candidate policies to consider. Email software could be programmed to use a beep with a delay, perhaps

[6] The precise turning point is the threshold that satisfies $1 = -\log(1 - p^*)$ which is $1 - 1/e$, roughly 0.63.

[7] Indeed $z = 1$ or beep-on is never optimal. Intuitively this follows from an envelope theorem argument. Consider a z very close to 1. Then if the researcher is lucky there will be no beep and she will work very long, call it $t(z)$ before checking. This however has low probability and the average waiting time puts most of the weight on stopping due to a beep. Now when we reduce z marginally, the researcher's optimal stopping rule is unchanged. She stops when there's a beep or when there is no beep before $t(z)$. So the effect on average stopping time is due to the direct effect of the shift in total probabilities of the two scenarios. A reduction in z shifts weight toward the preferred no-beep scenario. It also increases the expected time before a beep which further adds to the expected working time.

a random delay, perhaps a random delay that depends on intricate details of the prior and future history.

To characterize the optimal mechanism it would be intractable to optimize over the large set of feasible information policies. Fortunately, building on ideas from Kamenica and Gentzkow (2011), Aumann and Maschler (1995), Ely, Frankel and Kamenica (2013), we can appeal to the *obfuscation principle* and capture the full set of feasible policies by a tractable family of simple, *direct obfuscation mechanisms*[8]. Here is the logic. The principal's payoff depends on the agent's sequence of actions which in turn depend on the realized path of the agent's beliefs. Any information policy induces a stochastic process for those beliefs. However the process necessarily satisfies two constraints. First, by the law of total probability, the updated belief ν_t of the agent after observing a message at time t must be distributed in such a way that its expectation equals the belief μ_t held before observing the message. Second, after updating based on the message, the agent's belief evolves autonomously with the passing of time because the agent understands the underlying probability law, in this case the arrival process of email. For example, if the principal sends a message that leads the agent to assign probability ν_t to the presence of an email, the passage of time will cause this belief to trend upward because the agent is aware that email is arriving stochastically even if he doesn't directly observe its arrival.

We can express these properties as follows.

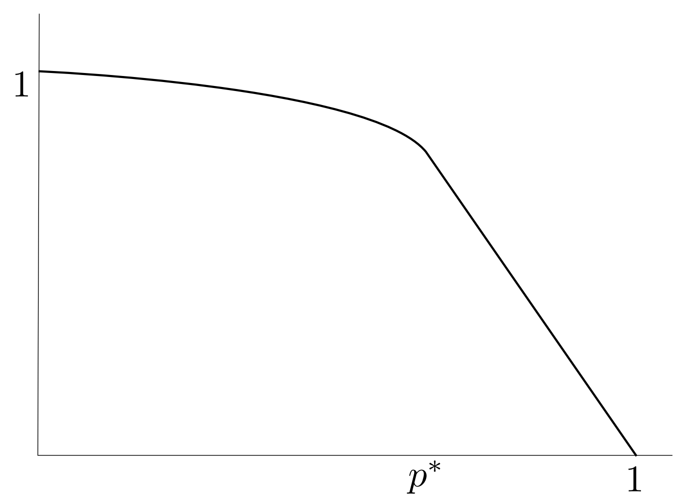
1. $\mathbf{E}(\nu_t \mid \mu_t) = \mu_t$,
2. $\dot{\nu}_t = \lambda\,(1 - \nu_t)$.

The obfuscation principle, proven for the general model below, asserts that in fact given any stochastic process for the agent's beliefs satisfying the two conditions there is an information policy that induces it. Indeed to find such a policy it is enough to suppose that the principal tells the agent directly what his beliefs should be after every history. As long as the sequence of recommendations follows a probability law that verifies the conditions above, the agent will always rationally accept the principal's suggested belief.[9]

[8] See also Sorin (2002) and Cardaliaguet et al. (2013).

[9] Similar to the revelation principle, these direct obfuscation mechanisms represent a

This enables us to reformulate the problem and solve it using dynamic programming, using μ_t as a state variable and explicitly incorporating the two constraints. I show for the general model below how to characterize the principal's value function using a series of operations that have a tractable geometric representation, building on concavification arguments in Kamenica and Gentzkow (2011) and Aumann and Maschler (1995). For the email beeps problem the optimal value function I obtain is depicted below.



Having derived the value function I then show how to infer the principal's optimal policy. To begin with, recall that in continuous time the constant arrival rate of email implies that the agent's beliefs are drifting monotonically upward to 1. Initially, when the agent's beliefs are below p^*, the value function shows that the principal's continuation value is decaying exponentially as this occurs. We can infer that the optimal policy for the principal is beep-off to the left of p^*. The principal enjoys the benefits of the agent working and allows the agent's beliefs to drift upward until the agent is just on the verge of abandoning work and checking email.

On the other hand, when μ_t is to the right of p^*, the linearity of the value function tells us that the optimal policy is a random beep assigning the agent two possible interim beliefs, namely $\nu_t = p^*$ and $\nu_t = 1$. This

canonical way to induce a chosen stochastic process but often the next step is to find a natural, intuitive, or practical *indirect* mechanism that also implements it. For the beeps problem we will indeed find an attractive indirect implementation.

gives the principal the weighted average value associated with those two beliefs with the weights being defined by the requirement in item 1 that the average belief equal μ_t, hence the linearity over this interval. In this region the principal is resigned to the fact that the agent cannot be dissuaded from checking email with probability 1. In light of this a random beep is chosen to maximize the probability of false negatives which induce the agent to continue working.

Each of the above are familiar strategies of information management from the static analysis in Kamenica and Gentzkow (2011). Roughly speaking when the agent's current beliefs induce him to choose the preferred action without intervention, do nothing. When intervention is required, maximize the probability of moving the agent back into the desired region. Neither in the beep-on region, nor the random beep region does the optimal policy need to make special use of the dynamic nature of the problem.

However, it's when the agent is right at the threshold p^* that the dynamics play a crucial role.[10] Beep-off is no longer optimal because the agent's beliefs will cross p^* and he will check with probability 1. However, a random beep is not optimal either. With beliefs exactly equaling p^*, the only feasible lottery over interim beliefs p^* and 1 is a degenerate lottery assigning probability 1 to p^*. But a degenerate lottery is also equivalent to beep-off. Thus, the principal must be doing something qualitatively different when the agent is right at p^*. Indeed, I show that the unique demands at p^* in fact give rise to a simple history-dependent optimal policy that can be applied globally, i.e. not just when the agent reaches p^*.

Consider a beep with a deterministic delay.[11] We will set the length of the delay, t^*, to solve

$$1 - e^{-\lambda t^*} = p^*.$$

In particular, t^*, is the time it takes for the agent's beliefs to reach $p^*$when the beep is off. Let's track the agent's beliefs when the beep is programmed to sound after a delay of length t^*. Starting at $\mu_0 = 0$ the belief begins trending upward. If an email arrives the beep will only sound t^* moments

[10]And note that under the dynamics induced by the optimal policy, beliefs will hit the point p^* an unbounded number of times. Indeed once μ_t passes p^* the principal will repeatedly send the agent back to p^* with positive probability.

[11]Toomas Hinnosaar first suggested this implementation. There are others, discussed in detail in Section 3 below.

later and therefore it's as if the beep-off policy is in effect for the initial phase up to time t^*. By construction, at the end of the initial phase the agent assigns exactly p^* probability to the presence of an email, just low enough to keep him working.

Now consider what happens in the very next instant. If the beep sounds it indicates that an email has arrived (at time zero), the belief jumps to 1, and the agent checks. Suppose instead that when the agent is at the threshold p^*, no beep is heard. He learns that an email has not arrived at any time equal to or earlier than t^* moments ago, and he learns nothing more than that. In particular the agent obtains no information about arrivals in the immediately preceding t^*-length time period. It is as if he has been under the beep-off protocol during that period. By construction, knowing for sure that there was no email t^* ago and applying beep-off since then keeps the beliefs pinned at p^* for as long as this state of affairs continues, i.e. until the first beep. Thus, the principal keeps the agent from checking for a total expected length of time equal to

$$t^* + 1/\lambda$$

since $1/\lambda$ is the expected time before the first beep once the agent has reached belief p^*.

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. In the main part of this paper I extend the analysis above 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 derive a geometric characterization of the solution and show in general that it can be derived by iterating a series of simple graphical operations. I then apply it to several examples to illustrate.

I also consider a few significant extensions. First I examine a simple example involving two strategically interacting agents. The agents are depositors at a bank which is at risk of default. The bank releases public and private information to the agents over time in order to prevent them from coordinating a run on deposits. In such an environment information disclosures control not just the agents' beliefs about the underlying state (here the health of the bank) but also their higher-order beliefs. A depositor's incentive to withdraw is determined not just by the likelihood of default, but also the likelihood that the other depositor is already running. Thus, in addition to managing the depositors' pessimism about default, the bank will try to manage each depositor's beliefs about the other's pessimism. I analyze this problem and show how the bank optimally uses private and minimally correlated disclosures to achieve this.

Next, I consider the extent to which the principal can further incentivize a patient and strategic agent by offering rewards or punishments in the form of more or less informative disclosures in the future. For example we may ask whether the principal can induce the agent to take the principal's preferred action in the short run in exchange for better information further in the future. To address these questions I analyze the dynamic principal-agent problem with a patient agent and characterize the optimal mechanism. The characterization is an extension of the geometric representation from the original model. I apply this characterization to the beeps example and show that in fact the delayed beep remains optimal even within the larger set of feasible mechanisms. Indeed I show that the mechanism holds the agent to his reservation value, namely the value he would obtain if he remained uninformed forever implying that the principal is capturing all the value from the information he discloses.

Finally, while most of the analysis in the paper is carried out in the context of a discrete-time version of the model, I also analyze the model directly in continuous time. The continuous-time analog of the Bellman equation, the HJB equation, has a similar geometric interpretation and can simplify the analysis of some problems. I use these results to analyze a 2-state, 3-action example which has some novel features.

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 same model as the baseline model in this paper. Specifically they consider a single agent who chooses among two actions and has short-run incentives. A principal with state-independent preferences designs an optimal policy for disclosing information about a hidden stochastic process. Although we study the same problem, our focus is different. 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. With two actions the optimal policy is quite often greedy but they show an example with three states in which it is not. Relatedly, I show with a simple example below that even with two states the optimal policy is typically not greedy if the restriction to two actions is removed. Beyond this I also study a principal with state-dependent preferences and an agent with a general action space. I also show how to extend the analytical techniques to the case of a strategic agent, I analyze an example with multiple agents who interact over time, and I consider both discrete-time and continuous-time approaches to the problem as each of these are convenient for different purposes.

There is a growing literature studying dynamic information disclosure policies to which the results in this paper have the potential to contribute. Che and Hörner (2013) and Kremer, Mansour and Perry (2013) both study a social experimentation problem in which agents decide in sequence whether to experiment with a new technology. The principal internalizes a positive externality from experimentation and decides how much information to disclose about past experiments in order to boost private incentives to take risks. Halac, Kartik and Liu (2014) consider a research contest in which the prospects for a successful innovation are uncertain. They study the optimal combination of monetary incentives and information revelation policies when the principal wants to motivate contestants to engage in research, see also Bimpikis and Drakopoulos (2014). Gershkov and Szentes (2009) study information aggregation in organizations in which information acquisition is private and costly. A principal decides which agents should incur the costs of information acquisition and what those agents should be told about the previous efforts of other agents.

I consider an extension in which the principal optimally designs intertemporal incentives for an agent who may induced to strategically distort his current behavior in pursuit of promised rewards or to avoid threatened punishments. The model is thus a dynamic principal agent relationship without transfers. Such problems have been studied in the literature, see for example the recent paper by Guo and Hörner (2014). However in my model the agent has control over all payoff relevant actions and the principal has only information as an incentive instrument. I show how the optimal mechanism can be characterized by geometric means similar to, but somewhat more complicated than, the baseline model.

Bayesian persuasion models have some counterparts in the setting in which the principal mediates between many strategically interacting agents. 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). I study a multi-agent dynamic persuasion model in the context of a bank run. The principal is a bank who optimally discloses public and private information about solvency to depositors. The principal's goal is to minimize the chance that the agents coordinate on a bank run.

2 Model

A principal privately observes the evolution of a stochastic process. An agent knows the law of the process but does not directly observe its realizations. She continuously updates her beliefs about the state of the process and takes actions. The principal has preferences over the actions of the agent and continuously sends messages to the agent in order to influence her beliefs and hence her actions. The principal commits to an information policy and the agent's knowledge of the policy shapes her interpretation of the messages.

Formally, there is a finite set of states S and the principal and agent share a prior distribution over states given by μ_0. State transitions occur in continuous time and 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. Absent any information from the principal, the agent's beliefs will evolve in continuous time according to the law of motion

$$\dot{\mu}_t = \lambda \left(M_\mu - \mu \right)$$

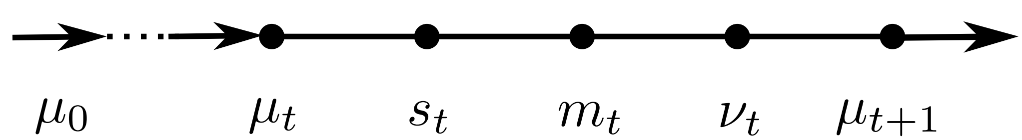
where $M_\mu = \sum_s \mu(s) M_s$.

We will begin with an analysis in discrete time and obtain continuous time results by on the one hand taking limits as the period length shortens, and on the other hand solving the dynamic optimization directly in continuous time. In discrete time with a given period length, this process gives rise to a Markov chain with transition matrix $\tilde{M}$ and a law of motion for beliefs given by

$$\mu_{t+1} = \mu_t \cdot \tilde{M}$$

which for notational convenience we represent by the (linear) map $\mu_{t+1} = f(\mu_t)$.

The principal sends messages to the agent in order to influence the evolution of beliefs. The timing is illustrated below.



The agent begins each period t with a posterior belief μ_t. Then the principal observes the current state s_t and selects a message m_t to send to the agent from the set of messages M_t. Next the agent updates to an interim belief ν_t and takes an action. Finally, time passes before the next date $t+1$ and the agent, knowing that the process is evolving behind the scenes, further updates to the posterior μ_{t+1}. A key observation is that because the principal controls all information available to the agent, he always knows the posterior μ_t and hence the agent's posterior is a natural state variable to be used in a dynamic programming characterization of the optimal mechanism.

The agent selects an action a_t in order to maximize the expected value of a flow state-dependent utility function x:

$$a_t \in \operatorname{argmax}_a \mathbf{E}_{\nu_t} x(a, s).$$

The principal's payoff $u(a, s)$ depends on the state s as well as the agent's action and therefore indirectly depends on the agent's interim belief, i.e. $u(\nu, s) = u(a(\nu), s)$. Indeed we can take this indirect utility function to be the primitive of the model and avoid getting into details about the agent's action space and payoff function. Henceforth we will assume that

$$u : \Delta S \times S \to \mathbf{R}$$

is a bounded upper semi-continuous payoff function for the principal[12] and that the principal is maximizing the expected value of his long-run average payoff

$$(1-\delta)\sum_{t=0}^{\infty}\delta^t u(\nu_t, s_t). \tag{1}$$

Finally, the agent's interim belief ν_t is the Bayesian conditional probability distribution over states and hence conditional on the agent having belief ν_t, the principal's expected payoff is

$$u(\nu_t) := \sum_{s\in S}\nu_t(s)u(\nu_t, s).$$

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 space of policies is unwieldy for purposes of optimization. The following lemma allows us to reformulate the problem into an equivalent

[12]When the agent is maximizing the expected value of x, there will be interim beliefs at which the agent is indifferent among multiple actions. As is standard, when we optimize the principal's payoff we will assume that these ties are broken in such a way as to render the principal's optimal value well-defined. This is captured in reduced-form by upper semi-continuity of u. In particular if the principal can approach a payoff be a converging sequence of interim beliefs, then he can in fact secure at least that payoff by implementing the limit belief.

one in which instead of choosing a policy, the principal is directly specifying a stochastic process for the agent's beliefs.[13]

Lemma 1 (The Obfuscation Principle)**.** *Any policy σ induces a stochastic process (μ_t, s_t, ν_t) satisfying*

1. $\mathbf{E}(\nu_t \mid \mu_t) = \mu_t$,
2. $\text{Prob}(s \mid \mu_t, \nu_t) = \nu_t(s)$ *for all* $s \in S$,
3. $\mu_{t+1} = f(\nu_t)$.

The principal's expected payoff from such a policy is

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

where the expectation is taken with respect to the stochastic process ν_t.

Conversely any stochastic process (μ_t, s_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. $\sigma(h_t) = \sigma(\mu_t, s_t)$.

The familiar intuition was given in the introduction. The proof, which has to contend with potentially infinite message spaces and histories, proceeds somewhat indirectly and is in Appendix A. There is a subtlety which is worth expanding upon. First of all, notice that the principal's objective function can be expressed entirely in terms of the beliefs of the agent, and the constraint set can be reduced to a choice of stochastic process for those beliefs. As a result the underlying state s_t of the stochastic process plays no role in the optimization. In particular we can treat the principal's continuation value at date t as if it depends only on the current beliefs μ_t and not on the current state s_t. This may be surprising because if the policy affects the agent's beliefs, it must be s_t-dependent. Two distinct current states imply distinct distributions over subsequent states, and therefore distinct continuation policies for the principal. These continuations can

[13] 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.

indeed affect long-run payoffs and therefore generate s_t-dependent optimal messages in the current period. However, because the goal is obfuscation and the principal has commitment power he refuses to respond to state-dependent long-run incentives.

With these preliminaries in hand, we can now solve the principal's optimization problem. Formally, the principal chooses a stochastic process for (μ_t, ν_t) satisfying item 1 and item 3 and he earns the expectation of Equation 1 calculated with respect to the chosen process. He chooses the process to maximize that expectation. As we argued previously, μ_t is a natural state variable for a dynamic programming approach to optimization. When the agent enters period t with belief μ_t, the principal informs the agent what his interim belief ν_t should be and the principal earns the flow payoff $u(\nu_t)$. Then the agent updates to a posterior $\mu_{t+1} = f(\nu_t)$ and the principal earns the associated discounted optimal continuation value. The Bellman equation is as follows.

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

Following Kamenica and Gentzkow (2011) and Aumann and Maschler (1995), the particular form of the constraint set ($\mathbf{E}\nu_t = \mu_t$) implies that the right-hand side maximization, and therefore the value function, can be expressed geometrically as the concavification of the function in brackets. The concavification is the pointwise smallest concave function which is pointwise no smaller than the function being concavified.[14] We obtain the following functional equation.

$$V = \operatorname{cav}\left[(1-\delta) u + \delta \left(V \circ f \right) \right]. \tag{2}$$

The novelty that arises in a dynamic optimization is that the value function itself enters into the bracketed formula. 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

[14]Geometrically, one can identify a function with its hypograph. Then the concavification is the function whose hypograph is the convex hull of the original hypograph. See Rockafellar (1997) who uses the terminology of the convex hull of a function. Also, see Section D.1 below where I define and characterize the concavification of a correspondence.

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 Appendix A.

Theorem 1. *The optimal value function is characterized by the functional equation in Equation 2. 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 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, below we solve several examples with just a series of diagrams.

2.1 Beeps

In discrete time the email beeps example can be described as follows. The set of states is $S = \{0, 1\}$ indicating whether or not an email has arrived to the inbox. The discrete time transition probability from state 0 to 1 is the probability within a period of length Δ that at least one email arrives and is given by $M = 1 - e^{-\lambda\Delta}$ yielding the following law of motion for the agent's beliefs when uninformed:

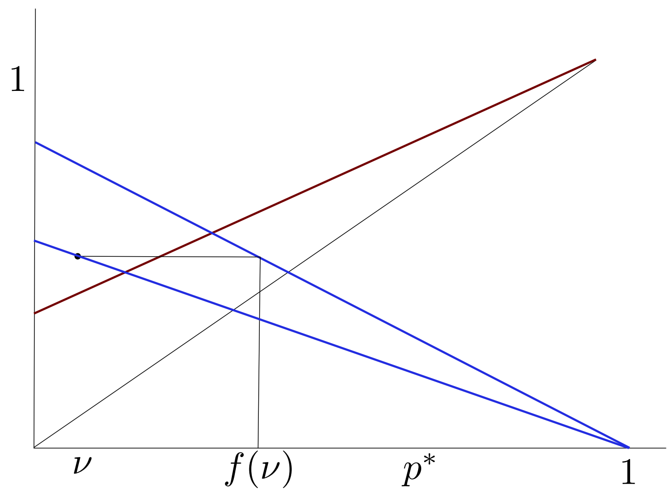
$$f(\nu_t) = \nu_t + (1 - \nu_t)M$$

The principal's indirect utility function is

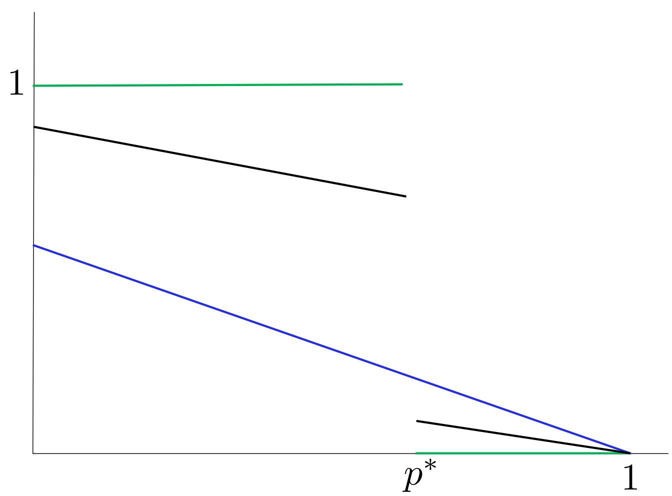
$$u(\nu) = \begin{cases} 1 & \text{if } \nu \leq p^* \\ 0 & \text{otherwise} \end{cases}$$

and he maximizes his expected discounted utility where the discount factor is $e^{-r\Delta}$ given a continuous time discount rate r.

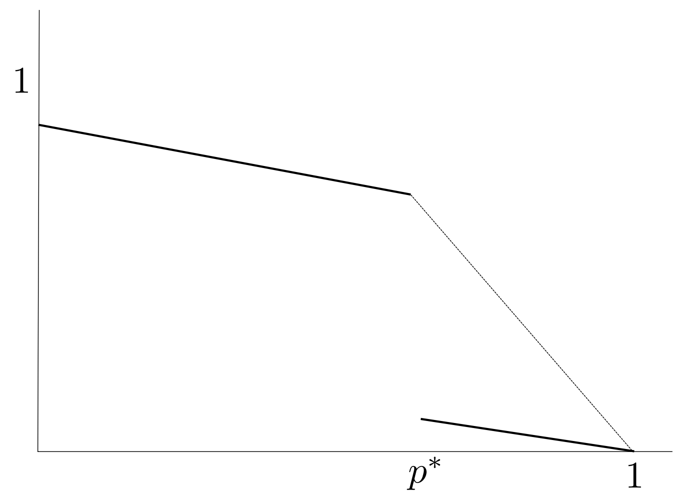
In Appendix A I derive the value function and optimal policy analytically. Here we will follow a sequence of diagrams to visualize the derivation and gain intuition. Refer to the Bellman equation in Equation 2. Consider as an initial guess, a linear V. We can trace through the operations on the right-hand side. The first step is to compose V with the transition mapping f. Since f is linear and $f(\nu) > \nu$, this composition has the effect of flattening V by rotating its graph to the left as illustrated in the following figure.



Next, we take the convex combination with the step function u yielding the discontinuous decreasing function below.



Lastly, we concavify the resulting function as illustrated in the next figure,



and we have the first iteration of the right-hand side operator. Notice that the function obtained differs from the initial candidate value function which is therefore not a fixed point and not the optimal value. In fact, since the beep-on mechanism discussed in the introduction yields a linear value function, we have shown that beep-on is not an optimal mechanism.[15]

Let us take stock of this first iteration and its implications for the optimal policy. Recall that the concavification represents the optimal lottery over interim beliefs, i.e. the optimal message distribution. At beliefs along a segment where the concavification differs from the underlying function, the optimal policy is to randomize between the beliefs at the endpoints of the segment. Thus in the interval $(p^*, 1]$, the principal wants to send the agent to either $\mu = p^*$ or $\mu = 1$ with the appropriate probabilities. At beliefs along a segment where the concavifcation and the underlying function coincide it is optimal to send no message, as is here between $\mu = 0$ and $\mu = p^*$.

The kink at p^* is a remnant of the discontinuity in the flow payoff u. It is easy to see that this kink will re-appear at every step of the iteration, as well as the linear segment from p^* to 1. What subsequent iterations add are additional kinks, first at the point $f^{-1}(p^*)$ in the second iteration, then

[15]We did not discuss how to interpret beep-on when the agent begins with a prior greater than zero. A fitting story is the following. The agent arrives to his office in the morning with a belief μ_0 that there is an email already there waiting for him. If indeed there is an email it will beep when his computer boots up, i.e. with probability μ_0. If it does not, then his belief jumps to 0 and beep-on continues from there. Thus, the value at μ_0 is just $(1 - \mu_0)V(0)$.

at $f^{-2}(p^*)$, etc. This occurs when we compose the drift mapping f with the previous iteration, shifting the kinks successively leftward. As we continue to iterate these are the qualitative features of the fixed point to which we converge.[16] The optimal value function is represented in Figure 1.

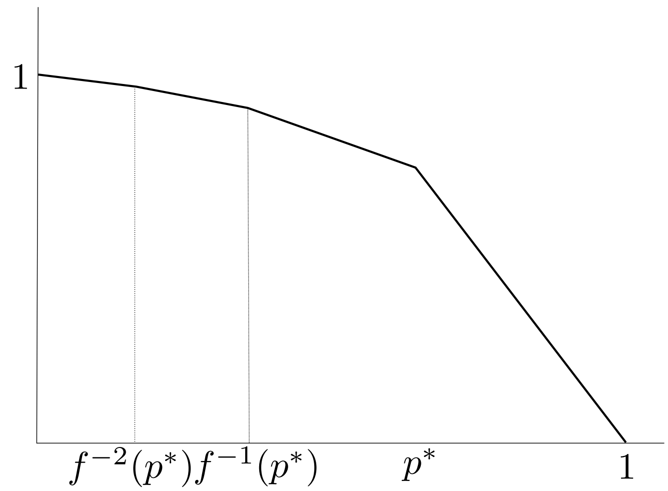


Figure 1: The optimal value function for the Beeps problem.

Now consider what happens to the optimal value function when we shorten the period length. In the shorter time interval the belief moves less between periods and f approaches the identity mapping. This has two implications. First, the number of kinks multiplies and in the limit the value function is smooth to the left of p^*. Second, the slope of the linear segment just to the left of p^* approaches the slope of the linear segment from p^* to 1. In the limit therefore, the value function is differentiable at p^* and indeed at every belief, see Figure 2.

It follows that to the left of p^*, it is uniquely optimal to send no message. In the discrete-time approximation, the linear segments between kinks allowed for a multiplicity of optimal policies ranging from randomization across the whole segment to no message at all. In continuous time, the strict concavity implies that any non-degenerate lottery is suboptimal. To the right of p^*, it remains optimal to randomize between p^* and 1. We have already discussed in the introduction that a delayed beep is optimal

[16] The number of kinks will be the maximum index k such that $f^k(0) \leq p^*$.

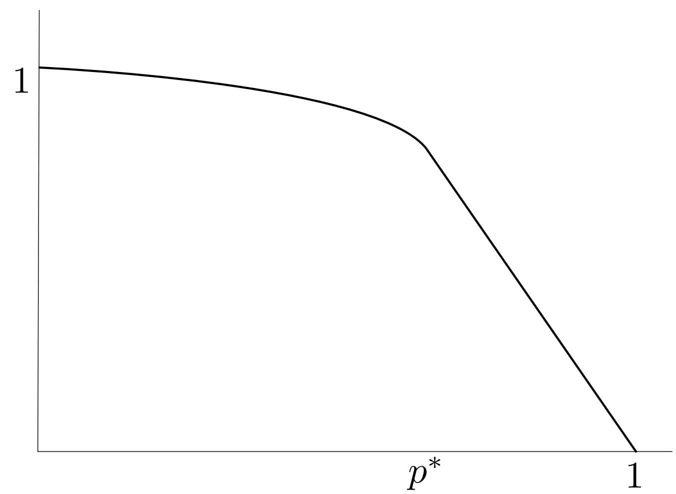


Figure 2: The continuous time value function.

when the beliefs are exactly p^*, and now we can elaborate further.

We can use the differentiability at p^* to compute the limiting value

$$V^*(p^*) \equiv \lim_{\Delta \to 0} V(p^*).$$

Indeed, at beliefs μ to the left of p^*, the value is given by

$$V^*(\mu) = \int_0^{t(\mu)} e^{-rt} dt + e^{-rt(\mu)} V^*(p^*)$$

where $\mu + (1-\mu)(1-e^{-\lambda t(\mu)}) = p^*$. That is the principal collects a flow payoff of 1 for a duration of $t(\mu)$ after which his belief reaches p^* whereupon his continuation value is $V^*(p^*)$. When we differentiate this expression with respect to μ and evaluate it at p^* we obtain the left-derivative of $V^*(p^*)$. The right derivative is simply the slope of the linear segment which is

$$\frac{-V^*(p^*)}{1-p^*}.$$

Using the fact that these one-sided derivatives are equal we can solve for $V^*(p^*)$ and we obtain

$$V^*(p^*) = r/(r+\lambda).$$

Notice that $r/(r+\lambda)$ is the discounted average value of receiving a flow payoff of 1 until a termination date which arrives at Poisson rate λ. Indeed that is exactly the *initial* (i.e. starting at $\mu = 0$, not at $\mu = p^*$) discounted value from the beep-on policy. In the optimal mechanism the principal obtains this value when the agent's beliefs are already at p^*, i.e. when t^* time has already passed during which he has been collecting a payoff of 1. Clearly this is accomplished by a beep of delay t^*. Moreover we can quantify the profit to the principal from using the optimal policy rather than beep-on. He is afforded an additional certain t^*-length duration in which the agent is working.

The results from this section are summarized as follows.

Theorem 2. *The optimal value function for the beeps example is*

$$V(\mu) = (1-\delta)\left[\sum_{s=0}^{n(\mu)-1} \delta^s + \delta^{n(\mu)}\left(\frac{1-f^{n(\mu)}(\mu)}{1-p^*}\left(1+\frac{\delta(1-p^{**})}{1-p^*-\delta(1-p^{**})}\right)\right)\right]$$

where

$$n(\mu) = \min\{n \geq 0 : f^n(\mu) > p^*\}.$$

In the continuous-time limit as the length of a period shrinks to zero, the value function converges to

$$\lim_{\Delta\to 0} V(\mu) = \left(1 - e^{-r(t^*-\tau)}\right) + e^{-r(t^*-\tau)}\left(\frac{r}{r+\lambda}\right).$$

for all $\tau \leq t^*$. This continuous-time optimum is implemented by a deterministic mechanism which notifies the agent after a delay of t^* after the first arrival.

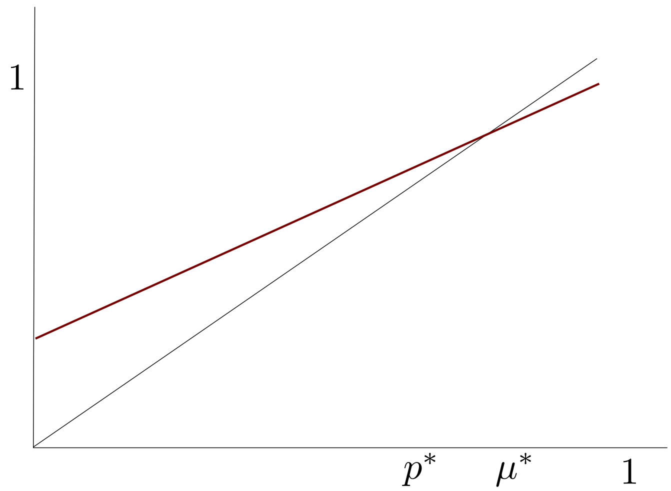
2.2 Ergodic Process

In the beeps example the state $s = 1$ is absorbing. When on the other hand the process is ergodic a new issue must be addressed. If the agent's belief reaches $\mu = 1$, it will begin to drift back to the interior, enabling further information revelation by the principal. How does the principal optimally incorporate this possibility?

To address this, consider now full-support transition probabilities so that the process admits a unique invariant distribution μ^* and let's assume

$\mu^* > p^*$.[17]

With μ^* as the invariant distribution, the law of motion for beliefs is no longer monotonic. Beliefs greater than μ^* move downward and beliefs below μ^* move upward. The mapping f crosses the 45-degree line at μ^*:



To aid the analysis, it helps to make a general observation about *absorbing sets* of beliefs. Say that an interval $\mathcal{I} \subset [0,1]$ is *absorbing under f* if $f(\mu) \in \mathcal{I}$ for all $\mu \in \mathcal{I}$. According to the following lemma, if u is linear over an interval that is absorbing under f, then the value function must also be linear over that interval.

Lemma 2. *Suppose that $\mathcal{I}$ is absorbing under f, and that u is linear over $\mathcal{I}$. Then V is also linear on $\mathcal{I}$.*

Proof. Consider any candidate value function W which is linear over $\mathcal{I}$. Since f is linear and u is linear over $\mathcal{I}$, the formula

$$(1-\delta)u + \delta\,(W \circ f) \tag{3}$$

must be linear over $\mathcal{I}$ because it is the convex combination of two functions which are linear over that interval. (That the composition is linear follows from the assumption that $\mathcal{I}$ is absorbing and W is linear on $\mathcal{I}$.)

[17] If $\mu^* \leq p^*$ the problem is simpler and less interesting. Eventually the beliefs will reach p^*. Once there the principal can cease sending messages and the agent will remain at p^* and work forever. The only problem to solve is how to get the agent to start working as quickly as possible when he begins away from p^*. It can be shown that this is accomplished by following exactly the strategy from the original beeps example.

The concavification of Equation 3 must be linear over $\mathcal{I}$. Thus, iteration starting with W must always stay within the set of functions which are linear over $\mathcal{I}$, i.e. the set of such functions is invariant under the value. Since the value mapping is a contraction, iteration converges globally to a fixed point, the fixed point must belong to any invariant set of functions. □

When $\mu^* > p^*$ the interval$(p^*, 1)$ is absorbing under f. Therefore the value function is linear there. It follows that the value function has the same shape as in the original problem. The optimal policy is therefore identical. In terms of implementation there is only one novelty: at $\mu = 1$ beliefs are now trending downward, i.e. the agent knows that eventually there will be a transition from state 1 to state 0. According to the optimal policy, the instant beliefs move into the interior the principal is randomizing between the two endpoints p^* and 1. This is achieved by a random message that reveals state changes but with false positives. As soon as a transition occurs the message is sent, but also each instant a transition does not occur the message is still sent but with a probability less than 1 calibrated so that the message induces interim belief p^*. Since the message has false positives but not false negatives, the agent remains at $\mu = 1$ as long as no message is heard.

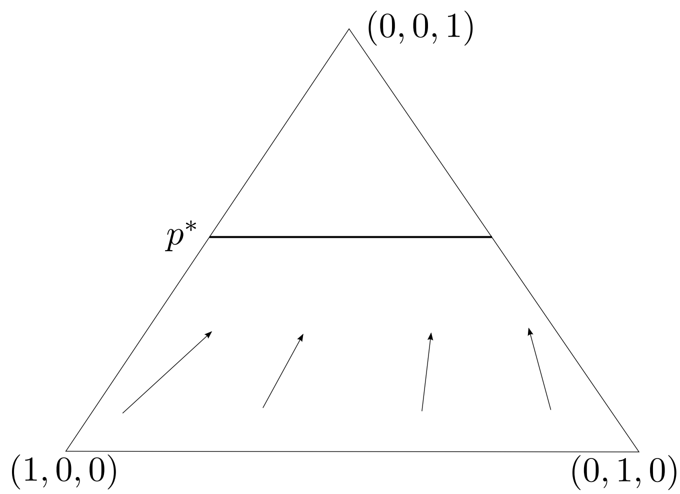
In particular, the value $V(1)$ is positive now because the agent will periodically switch back to working when his beliefs jump down to p^*.

2.3 Three States

When there are three states, $S = \{0, 1, 2\}$ and two actions, the threshold is no longer a point but a line segment through the simplex of beliefs ΔS. One one side of the line the agent takes action 0 and on the other side he takes action 1. The belief dynamics operate in a 2-dimensional simplex and can therefore be significantly more complicated. In this section I analyze a simple extension of the beeps problem to 3 states to illustrate.[18]

[18]This example is special because the belief dynamics have a monotonicity property. Once the beliefs leave the region where the agent takes a given action they never return (absent information from the principal). With two states the belief dynamics always have this property. With three states the beliefs may cycle and move up and down the steps of the principal's payoff function u on their path toward the invariant distribution. Such problems are considerably more complex to solve. In independent work Renault, Solan

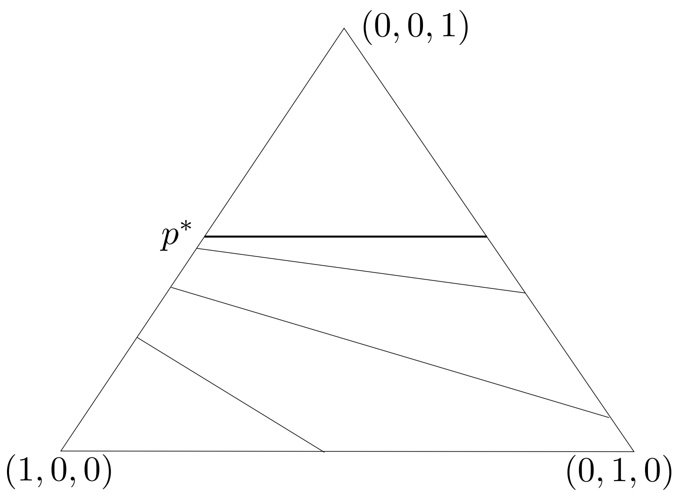
Consider a variation of the email problem in which the agent wishes to check as soon as the probability is sufficiently large, at least p^*, that there are at least *two* emails to read. Let $s \in S = \{0, 1, 2\}$ denote the number of unread emails currently in the inbox, where $s = 0$ and $s = 1$ indicate the exact number and $s = 2$ indicates 2 or more. Maintain the assumption that email arrives at rate λ. The following diagram illustrates the situation. The simplex is the set of possible beliefs for the agent and the line depicts the threshold above which the agent checks. The arrows show the flow of the agent's beliefs when the principal withholds information. The constant arrival of email implies that the beliefs move up and to the right.



This analysis of this problem follows similar lines as the two-state email beep problem. In the diagram below the lines represent points which lead by iterations of f to the threshold line. If we begin with a candidate value function which is linear and equal to zero at the $s = 2$ vertex, iterations lead to a piecewise linear value function with kinks along these segments.[19] The continuous time limit value function will therefore be linear along these line segments but strictly concave along rays toward the $s = 2$ vertex in the region below the threshold. It will be linear above the threshold and equal to zero at the $s = 2$ vertex. Thus, the optimal mechanism is a delayed beep signaling a past arrival of the second email.

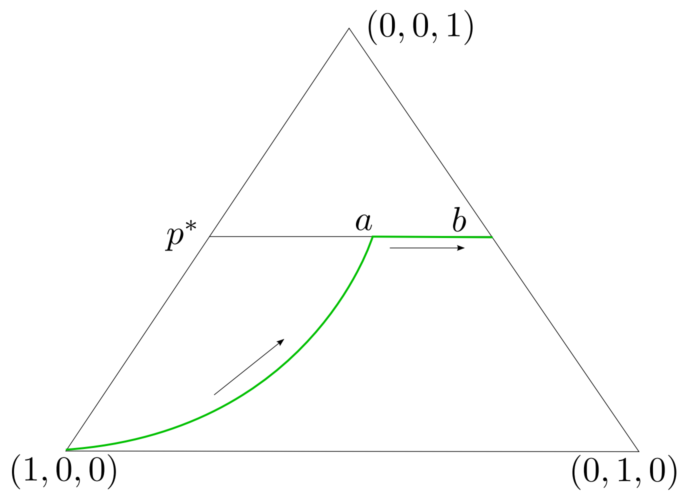
and Vielle (2014) analyze the many-state problem in more detail.

[19]These are not the level sets. As we will show below, the value declines as we move to the right along the p^* line. This pattern will thus be preserved at all of its inverse images under f.



The novelty that arises with three states concerns the evolution of beliefs and continuation values along the threshold. At the threshold the principal's policy is designed to maximize the probability that the agent continues working. As usual this is accomplished by sending the agent either to the threshold or to the most distant point in the shirk region (with the appropriate probability), in this case the vertex $(0, 0, 1)$ where the agent is certain that two emails are waiting. This signal allows the agent to increase his belief that a single email has arrived and thus as long as the agent remains at the threshold, this belief will continue to trend upward, converging toward the right face of the simplex. Note that the right face, where the agent is certain that at least 1 email is waiting, is isomorphic to the original 1-dimensional beeps problem because the agent is simply waiting to find out if one more email arrives.

The following diagram shows the path of the agent's beliefs. The beliefs will follow this path until they reach the threshold, then remain on the path until a beep sounds. As long as there is no beep the beliefs will converge asymptotically to the right face.



It follows that the length of the optimal delay must change as time passes. To see this, first consider the delay length at the point a where the beliefs first touch the threshold. Let's determine the delay length that keeps the agent on the threshold. Let t_a denote the length of time it takes for beliefs to reach t_a from the vertex $(1, 0, 0)$. If the beep has delay t_a then when the agent is at a and hears no beep his updated beliefs continue to attach probability p^* to the presence of 2 emails. The absence of a beep tells the agent that a second email did not arrive t_a moments ago. The key question is what does this information tell the agent about the conditional distribution over the remaining states $s = 0$, $s = 1$. At that point in the past he was at the point $(1, 0, 0)$. In particular he is certain that not even the first email had arrived. Learning that a second email did not arrive at that moment gives him no information about the other states since the simultaneous arrival of two emails has probability zero.

Thus, the absence of a beep tells him that his beliefs at the time t_a ago were correct, and thus that his current beliefs should be updated from those prior beliefs based only on the information that a time period of length t_a has passed during which he learns nothing about arrivals. By construction that updated belief assigns exactly p^* to $s = 2$.

By contrast, consider a point like b, further to the right. At this point he attaches higher probability to the presence of a single email. Suppose the principal continued to use a beep with delay t_a. What does the agent believe conditional on hearing no beep? He learns that as of t_a ago, the second email had not arrived. At that point in the past he assigned positive probability to the arrival of the first email. Of course the absence of the second email is information: it makes it less likely that the first email arrived. But nevertheless he will continue to assign positive probability to the event that a first email had arrived t_a ago. To obtain his current beliefs he will update that posterior based on knowing that t_a time has passed. His updated belief that a second has arrived during that time will be larger when starting from a positive probability of a first email than when starting with probability zero. The latter starting point would lead him to p^*, so using the delay length t_a would put him above p^*.

Therefore, in order to keep the agent on the threshold, the delay length must be shorter the more time has passed. In particular if the second email arrives at time t' then the delay before beeping must be shorter than if the second email arrived at time $t < t'$. What happens to this delay length asymptotically as time increases? Since the beliefs are approaching probability 1 that exactly 1 email is waiting, the length of time it takes for the probability of a second email to equal p^* converges to the length of time it would take if the agent were certain at the outset that the first email had already arrived. This is just the length of time for the arrival of a single email and that is the length t^* from the 1-dimensional problem.

We can understand in these terms why the continuation value must decline as we move toward the right face. Because the delay is shortening but the arrival rate of email is constant, it follows that beeps are arriving more quickly and thus the agent is jumping sooner on average to the upper vertex.

2.4 Three Actions

The solutions for each of the examples considered thus far are special in at least two senses. First, as was shown by Renault, Solan and Vieille (2014) the optimal mechanism is *greedy*; in particular it is nearly identical to the optimal mechanism in a static version of the problem. Second, the optimal policy is monotonic in that there is an initial phase of silence leading eventually to a phase of random messages. These features are typical of problems with two states and two actions. In this subsection I consider the simplest 3 action problem in which the optimal mechanism is very different than the static problem and in particular consists of an early phase of messages followed by a duration of silence and then ultimately a final message phase.[20]

Consider the following indirect utility function with 3 steps:

$$u(\nu) = \begin{cases} 5/4 & \text{if } \nu = 0 \\ 1 & \text{if } \nu \in (0, 1/2] \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

Here is a story that goes with it. When the agent is certain there is no email waiting he can work with full concentration. When there is a small chance of an email he is tempted to check but he resists the temptation. The effort spent on willpower makes him somewhat less productive. Finally when his beliefs cross the threshold (here $p^* = 1/2$) he succumbs to temptation and stops working.

The interval $(1/2, 1]$ is absorbing and u is linear there and so by Lemma 2 the value function has the familiar linear segment and the optimal mechanism randomizes between $1/2$ and 1 when the beliefs are anywhere in between. To analyze the interval $[0, 1/2)$ note that one feasible mechanism for the principal is to use the optimal mechanism from the basic beep example. That would yield a value function which is differentiable at the threshold $1/2$. Even though this is not optimal for the present problem, it places a lower bound on the optimal value function. In particular, the optimal value function cannot have a kink at $1/2$. As we will show below it is in fact smooth and strictly concave over some interval $(p^{**}, 1/2]$ so that no message is optimal there. However, unlike the original beeps problem, $p^{**} \neq 0$ and on the interval $[0, p^{**}]$ the value function is again linear. Intuitively, randomizing between 0 and p^{**} allows the principal to stay at the point 0, earning the high flow payoff of 1 for some duration, whereas following the original beeps solution the beliefs would spend only an instant there. (To complete the argument we must show that the gain from pausing at $\mu = 0$ for some time is not outweighed by the loss from having to move quickly to p^{**}. The formal argument below takes care of this point.)

As we will show below, this solution differs from the static solution on the interval $(p^{**}, 1/2)$. Another distinguishing feature of this example is that the optimal policy uses both false negatives and false positives. In the initial phase, the signals that move from 0 to p^{**} are false positives:

[20] Renault, Solan and Vieille (2014) also present an example in which the optimal mechanism is non-greedy involving two actions but three states. Two-state problems are generally easier to analyze so the example below is simpler.

conditional on the agent receiving this signal he changes his behavior but there is a probability $1 - p^{**}$ that this signal was received even if no email arrived. On the other hand, once the agent reaches $\mu = 1/2$ the optimal policy reverts to the false negatives from the original email problem.

3 Multiple Agents

Consider now the case of many agents. Whereas in the single-agent problem the agent's optimal action depended only on his belief about the state, in the multi-agent case each agent's optimal action will depend not only on his belief about the state but also his belief about the action of the other agent. It will therefore typically be in the interest of the principal to adopt a mechanism which selectively releases information about either or both. For illustration we will consider in this section the example of a bank run.

The bank may be either healthy or distressed. There are two agents and each has a deposit with the bank. Each agent wishes to withdraw his deposit, i.e. run, if the bank is distressed. Withdrawal is an irreversible decision. However, even if the bank is healthy each agent nevertheless wishes to run if he expects that the other agent will run. The bank can survive if only one agent withdraws. We will say that a *bank run* occurs only if both agents withdraw. In the case of a bank run, the bank is insolvent and suffers a loss. This has several implications. To begin with, notice that agent 1 will run if he considers it sufficiently likely that the bank is distressed *or* if he considers it sufficiently likely that 2 considers it sufficiently likely that the bank is distressed. And of course agent 2 will run if he thinks the bank is sufficiently likely to be distressed or that agent 1 is sufficiently likely to fit the description in the previous sentence. To summarize, the agent's optimal actions are no longer a function solely of their beliefs about the state, but nevertheless they can be expressed as functions of their beliefs and *higher-order* beliefs.[21]

To formalize this idea we will need some background definitions and notation. Let the set of states be $S = \{\texttt{healthy}, \texttt{distressed}\}$. The set of all infinite hierarchies of belief for agent i about S will be denoted H_i. An element $\beta_i \in H_i$ describes not only agent i's belief about S (these are called

[21] Angeletos, Hellwig and Pavan (2007) and Morris and Shin (1999) are two papers that study dynamic models of bank runs but with an exogenous flow of public and/or private information.

first-order beliefs), but also his belief about agent $-i$'s beliefs about S (the second-order beliefs) and indeed i's belief about the probabilistic relationship between S and the first-, second-, and any arbitrary kth-order belief of agent $-i$.

It is well-known[22] that each element β_i of H_i can be uniquely (and more compactly) represented as a single probability measure

$$\beta_i \in \Delta(S \times H_{-i}).$$

Messages sent by the principal to the two agents (possibly publicly, possibly privately) will affect the agents' beliefs and in turn their actions and thus the pair of beliefs (β_1, β_2) will act as the multi-agent counterpart of μ_t, acting as the state variables for the principal.

To model the agents' preferences, consider the payoff tables below, where $p^* \in (0,1)$. They show the payoffs an agent earns from either of the two actions $\{\texttt{wait}, \texttt{run}\}$ as a function of the state and of the action of the other agent. The payoffs of the two agents are symmetric.[23]

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

`healthy`

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

`distressed`

Note that `run` is the uniquely optimal action whenever one of the following is true: either the probability of state `distressed` is larger than p^* or the probability that the opponent plays `run` is larger than p^*. Indeed, `run` is the uniquely optimal action exactly when the total probability of the event "either the bank is distressed or the other agent plays `run`" is greater than p^*. This observation translates into a rule that gives the uniquely optimal[24] action as a function of the agent's belief β_i, as we now show.

[22]See Mertens and Zamir (1985), Brandenburger and Dekel (1993), or Ely and Peski (2004)

[23]These payoffs are special in that the single parameter p^* determines the best-responses. The example in this section is mean to illustrate how the techniques can be extended to multi-agent problems. A comprehensive analysis of general games and payoffs is left for future research.

[24]Formally, uniquely rationalizable.

Let ρ_i^k be the set of beliefs defined inductively as follows.

$$\rho_i^1 = \{\beta_i : \beta_i(\{\texttt{healthy}\} \times H_{-i}) < 1 - p^*\}$$
$$\rho_i^k = \left\{\beta_i : \beta_i(\{\texttt{healthy}\} \times \neg\rho_{-i}^{k-1}) < 1 - p^*\right\}$$

If agent i's belief is an element of the set ρ_i^1, then he believes with a probability greater than p^* that the bank is distressed and his uniquely optimal action is to run. If i's belief is an element of ρ_i^2 then the probability that *either* the bank is distressed or agent $-i$ will run (due to the fact that agent $-i$'s belief belongs to ρ_{-i}^1) is greater than p^* and again his uniquely optimal action is to run. By induction the same conclusion applies to beliefs belonging to all sets ρ_i^k. Now set

$$\rho_i = \cup_{k\geq 1}\rho_i^k$$

The unique rationalizable action for an agent with belief in ρ_i is `run`. This implies that no matter what (Bayesian Nash) equilibrium is played by the agents, agent i's strategy must choose the action `run` when his belief belongs to ρ_i. In fact, the best equilibrium for the bank is the one in which an agent runs exactly when his belief belongs to ρ_i.

Proposition 1. *Suppose that the agents' belief profile (β_1, β_2) is drawn from some joint probability measure $\pi \in \Delta(H_1 \times H_2)$ and the agents play a Bayesian Nash equilibrium of the associated game of incomplete information. Then the following symmetric strategy profile*

$$\alpha_i(\beta_i) = \begin{cases} \texttt{run}, & \text{if } \beta_i \in \rho_i \\ \texttt{wait} & \text{otherwise.} \end{cases}$$

is one such equilibrium and it is the equilibrium that minimizes the probability that either agent plays `run`.

If we assume that the bank's objective is to minimize the probability and/or delay the onset of a bank run, i.e. withdrawal by both agents, then we can model the bank's (state-independent) payoffs as follows.

Then it follows from Proposition 1 that regardless of the bank's mechanism for releasing information, the best equilibrium from the bank's perspective is the one identified above. This result allows us to analyze the

	wait	run
wait	0	0
run	0	-1

multi-agent mechanism design problem in the same way as we did in the single-agent case. In particular we can fix in advance the decision rule of the agents (as a function of their beliefs) and then focus the analysis on the design of a mechanism for manipulating beliefs.[25]

A formal statement of this idea appears below. It implies that, when studying the incentives of agent i, we can collapse the entire hierarchy of beliefs into a belief over two auxiliary "states", namely

$$\omega_i := \{\texttt{healthy}\} \times \neg\rho_{-i},$$

and that set's complement. In particular, we can treat i's belief β_i as a single probability in $[0, 1]$, the probability he assigns to the event ω_i.

Lemma 3. *Agent i's belief β_i belongs to ρ_i (and therefore he chooses* `run`*) if and only if*

$$\beta_i(\omega_i) < 1 - p^*$$

Mechanisms Let's suppose that when the process begins it is common knowledge that the bank is healthy but that the bank's state switches to `distressed` at Poisson rate λ. One interpretation is that at time zero there has been an exogenous public event which raises doubts about the bank's health, potentially causing a run. We are modeling the bank's incentives to disclose further information about its health in order to forestall the run.[26]

One mechanism the bank may use is public disclosure: release public signals about the current state of the bank. Public disclosure mechanisms can be analyzed using methods that are identical to the single-agent case

[25] We are following standard practice in mechanism design by selecting the equilibrium that is best for the principal. We can interpret this selection as arising from a procedure in which the principal, after sending an informative message, recommends an action to the agent. The strategy profile we have selected is then "incentive-compatible" and the best among all incentive-compatible mechanisms.

[26] Since the agents are playing a coordination game and coordination is facilitated by common belief about the state, we can think of the principal's problem as controlling the flow of information to minimize *common learning*, see Cripps et al. (2008).

because with public disclosure it will be common knowledge that each agent has identical beliefs about the state. In particular, take any feasible policy for the single-agent problem given by Lemma 1, and there exists a public disclosure mechanism in which it will be common knowledge that both agents' beliefs follow the same induced stochastic process, and conversely. Thus, among public disclosure mechanisms the problem of maximizing the bank's payoff is equivalent to maximizing the delay before a single agent's belief exceeds p^*. It follows from our previous results that the optimal public disclosure mechanism is a delayed public signal. When the state switches to `distressed`, the bank waits a length of time t^* before releasing a public signal which perfectly reveals that the state is `distressed`. The expected timebefore the onset of the bank run is given by

$$t^* + \frac{1}{\lambda}$$

Indeed, our results for the single-agent problem imply that $t^* + 1/\lambda$ is the maximum expected length of time the principal can either *individual* agent to switch to `run` under any mechanism, public or otherwise. This is because we showed that $t^* + 1/\lambda$ is the maximum expected length of time that a single agent's beliefs that the state is `distressed` can be held at or below p^*. In the multi-agent case, since agent i can be induced to play `wait` if and only if $\beta_i \notin \rho_i$, agent i switches to `run` as soon as he assigns greater than p^* probability to `distressed`, and this occurs no sooner than $t^* + 1/\lambda$ in expectation.

Say that a process for the individual belief β_i is *individually optimal* with respect to agent i if it achieves this upper bound, i.e. if the expected time before $\beta_i(\{\texttt{healthy}\} \times H_{-i}) > p^*$ is equal to $t^* + 1/\lambda$. A delayed beep mechanism for agent i is one example of an individually optimal process with respect to i. (There are many others as discussed below.) We can understand a public delay mechanism as achieving a joint process for the pair of beliefs (β_1, β_2) which is individually optimal with respect to each agent but which is perfectly correlated, i.e. $\beta_1 = \beta_2$ at all points in time.

The bank can do better by designing a joint process for (β_1, β_2) that was individually optimal for each agent but not perfectly correlated. Then, the time at which a bank run occurs will be

$$\max\{\tilde{r}_1, \tilde{r}_2\}$$

where $\tilde{r}_i$ is the random time at which agent i runs, i.e the first instant at which $\beta_i(\{\texttt{healthy}\} \times H_{-i}) > p^*$. As long as the process (β_1, β_2) is individually optimal (so that the expected value of $\tilde{r}_i$ is $t^* + 1/\lambda$) but less than perfectly correlated, the expected value of max $\{\tilde{r}_1, \tilde{r}_2\}$ will be larger than $t^* + 1/\lambda$, thus improving on public disclosure.

Indeed the smaller the dependence between $\tilde{r}_1$ and $\tilde{r}_2$, the larger will be the bank's payoff in expectation so that bank prefers private disclosure with minimal correlation. However, there is a limit to the degree to which the two belief processes β_1 and β_2 can be decoupled. Intuitively, the two agents' beliefs evolve in response to learning about the same underlying state, namely the health of the bank and this entails some correlation in beliefs. Formally, according to Lemma 3 if agent i attaches high probability to ω_i, i.e. $\beta_i(\omega_i) \geq 1 - p^*$, then in particular $\beta_i(\{\texttt{healthy}\} \times \neg\rho_{-i}) \geq 1 - p^*$ implying $\beta_i(\neg\rho_{-i}) \geq 1 - p^*$ or in other words $\beta_i(\{\beta_{-i}(\omega_{-i}) \geq 1 - p^*\}) \geq 1 - p^*$. That is, at any history in which i attaches high probability to ω_i, i also attaches high probability to $-i$ attaching high probability to ω_i. As an implication of this, conditional on the event that i attaches high probability to ω_i, an outside observer will attach high probability to $-i$ attaching high probability to ω_i. In other words, β_1 and β_2 are correlated.

While it is not possible to achieve arbitrary degrees of (un)correlation between $\tilde{r}_1$ and $\tilde{r}_2$, we now describe an alternative individually optimal mechanism which can be coupled with the delay mechanism to produce a joint process which improves upon public disclosure. To begin with, consider the following *lamppost* mechanism. Agent i is given no information prior to date t^*, i.e. up until he assigns exactly probability p^* to state `distressed`. From date t^* onward the mechanism works as follows. First, if the state was `healthy` at date t^*, then the mechanism is set to beep as soon as the state switches to `distressed`. On the other hand, if the state had already switched to `distressed` at date t^*, then an auxiliary Poisson clock is started with parameter λ and set to beep as soon as the clock goes off. All beeps are private and sent only to agent i.

The lamppost mechanism is individually optimal with respect to agent i. To see this note first that the distribution of beep times in the lamppost mechanism is identical to the distribution generated by a beep with an optimal delay. Under either mechanism the beep sounds at a random time $\tilde{r}_i \in [t^*, \infty]$ where the distribution is exponential with parameter λ. Next, note that the lamppost mechanism generates the same information for the agent as the optimal delayed beep. Prior to date t^* the agent is uninformed and his beliefs continuously increase toward p^*. Once the beliefs reach p^* (at date t^*), a beep occurs only if the state is `distressed` and thus a beep leads the agent to assign probability 1 to that state. Consider the agent's beliefs when, at some date $t \geq t^*$ no beep has occurred. First, the agent knows that no state change has occurred between t^* and t otherwise there would have been a beep. Second, the absence of a beep occurs with the same probability whether or not a state change had already happened prior to date t^*. Thus, the agent will never obtain information about the possibility of a state change from `healthy` to `distressed` at any date between 0 and t^*. By the definition of t^*, the Bayesian conditional probability of state `distressed` is exactly p^* when the agent has no information about state changes during a period of length t^* and knows for sure that no state change has occurred at any other time. This all shows that at every history of the process the beliefs induced by the lamppost mechanism are identical to the beliefs induced by the optimal delayed beep mechanism.

The bank can improve over a public delay mechanism by using two distinct individually optimal mechanisms and sending their messages privately to the two agents. To gain some intuition, we can analyze the dynamics of higher-order beliefs under such a mechanism. To begin with the delay mechanism and the lamppost mechanisms share the same dynamics for higher-order beliefs during the initial period up to time t^*. Under both of these mechanisms, with probability 1, each agent obtains no information and chooses `wait`. Thus, for all dates prior to t^*, it is common knowledge that each agent's first-order probability of state `distressed` is $1 - e^{-\lambda t}$. Thus, when date t^* is reached, it is common knowledge that each agent assigns probability p^* to state `distressed` and the complementary probability to state `healthy`. Moreover, it is common knowledge that neither agent has switched to `run`. To summarize, at date t^*, each agent has beliefs represented by the table in Figure 3 where agent's belief is described by a probability over the health of the bank (the columns) cross the hierarchy of beliefs of the other agent (the rows).

Consider what happens next. Because each agent is at the threshold p^* for first-order beliefs at which `run` becomes a dominant action, each agent knows that the other will begin switching to `run` at some rate. When the bank uses a public mechanism, each agent switches to `run` at precisely the same time. This means that as long as agent i has not switched, not only does his first-order probability of `distressed` remain at p^*, but also he continues to believe with probability 1 that $-i$ has not switched. In other

	healthy	distressed
$\neg\rho$	1 - p^*	p^*
ρ	0	0

Figure 3: At time t^*, each agent assigns probability p^* to `distressed`, and is certain that the other agent has beliefs in $\neg\rho$.

words, his belief is constant and equal to the belief in the table above.

By contrast, consider what happens after t^* when the bank uses a delayed beep mechanism for agent 1 and a lamppost mechanism for agent 2. First consider the dynamics of higher-order beliefs for agent 1. When we reach date $t^* + \Delta$, agent 1 knows that if there was a state change in the interval $[t^*, t^* + \Delta]$, agent 2 will have been informed of this (by the lamppost mechanism) and will have switched to `run`. As long as $\Delta \leq t^*$, this represents a period of time that is shorter than the length of delay in agent 1's mechanism. Thus agent 1 continues to choose `wait` despite assigning his increasing belief that agent 2 has switched to `run`. Agent 1 is willing to do this because the event that 2 has been informed of a state change and switched to `run` is a *subset* of the p^*-probability event that there was a state change that 1 has yet to be informed of. In other words, the p^* probability in the top-right corner of the table above is just being spread out across the right column. Agent 1 continues to assign exactly probability $1 - p^*$ to the event ω_i (i.e. the top-right corner). Indeed, over the time interval $[t^*, 2t^*]$, agent 1's higher-order beliefs evolve according to the table in Figure 4, where γ_t increases from 0 to 1 as t moves from t^* to $2t^*$.

	healthy	distressed
$\neg\rho$	1 - p^*	(1 - γ_t) p^*
ρ	0	$\gamma_t p^*$

Figure 4: Dynamics of higher-order beliefs from t^* onward.

Examining these dynamics we see how using the private pair of mechanisms allows the bank to capitalize on some slack in the higher-order beliefs. It allows agent 2 to switch to `run` without this causing agent 1 to switch as well. In particular, agent 1 will switch *after* agent 2. Since agent

2's mechanism is individually optimal, the time at which a bank run ensues must therefore be longer than $t^* + 1/\lambda$.

Finally, once we reach date $2t^*$, and forever thereafter, as long as 1 has not heard a beep, he knows that due to the delay of length t^* it is possible that there was a state change at some time during the previous t^*-length window, and moreover if that has happened, agent 1 knows that agent 2 must have learned about it and switched to `run`. Since this event always has probability p^*, agent 1's higher order beliefs will remain constant and equal to the belief represented in the following table.

	`healthy`	`distressed`
$\neg\rho$	1 - p^*	0
ρ	0	p^*

Figure 5: Agent 1's beliefs at time $2t^*$ and onward until he withdraws.

Examining these beliefs, we can see that they represent the limiting boundary of higher-order beliefs at which `wait` is an incentive-compatible action. This is because the probability of state `distressed` cannot decline and the probability that the opponent plays `run` also cannot decline. The pair of private individually optimal mechanisms pushes higher-order beliefs to this boundary and holds them there as long as possible. From this date forward, agent 1 either hears a beep and switches to `run` or hears no beep and remains at the limiting belief.

The dynamics for agent 2's higher-order beliefs under the lamppost mechanism are identical. The only difference is that the time window during which a state change may have occurred (and agent 1 learned about it and switched to `run`) is the initial $[0, t^*]$ period, as opposed to the most recent period of that length.

Nevertheless there remains one additional margin for improvement. Recall that the lamppost mechanism involves an auxiliary clock which is started at time t^* if the state had already changed. There is some freedom to design the correlation between the time $\psi \geq t^*$ at which this clock beeps and the time $\phi \leq t^*$ which the state changed. The former has an exponential distribution on $[t^*, \infty]$ and the latter has a (truncated) exponential distribution on $[0, t^*]$. Therefore since agent 1 has a delayed beep, he switches to `run` at $\tilde{r}_1 = \phi + t^*$, and conditional on $\phi \leq t^*$ agent 2 switches to `run` at

$\tilde{r}_2 = \psi$ so that a bank run occurs at the following time

$$\max\{\tilde{r}_1, \tilde{r}_2\} = \max\{\phi + t^*, \psi\}$$

with conditional expected value (conditional on the state change occurring prior to t^*)

$$\int_{s=0}^{\infty} 1 - G(s)ds \tag{5}$$

where

$$G(s) = \text{Prob}\{\tilde{r}_1 \leq s \text{ and } \tilde{r}_2 \leq s\}$$

is the conditional CDF of $\max\{\tilde{r}_1, \tilde{r}_2\}$.

From the theory of copulae in probability theory, and in particular Sklar's Theorem and the Fréchet-Hoeffding Theorem (see Nelsen (1999)) we can minimize G pointwise, and therefore maximize the expected delay length in Equation 5 by adopting as the the joint distribution for ϕ and ψ (whose marginal distributions we are given) the *Fréchet-Hoeffding lower bound*. In this case the lower bound is achieved by setting

$$\psi = -\frac{1}{\lambda}\log\left(\frac{F(\phi)}{p^*}\right) + t^* \tag{6}$$

for all $\phi \leq t^*$ where F is the (marginal) CDF of ϕ, (i.e. the CDF of the exponential distribution with parameter λ.) To verify, conditional on $\phi \leq$

t^* the distribution of $\psi - t^*$ is

$$
\begin{aligned}
\text{Prob}\left(\psi - t^* \leq t \mid \phi \leq t^*\right) &= \text{Prob}\left(-\frac{1}{\lambda}\log\left(\frac{F(\phi)}{p^*}\right) \leq t \mid \phi \leq t^*\right) \\
&= \text{Prob}\left(F(\phi) \geq p^* e^{-\lambda t} \mid \phi \leq t^*\right) \\
&= \frac{\text{Prob}\left(\left\{F(\phi) \geq p^* e^{-\lambda t}\right\} \cap \left\{\phi \leq t^*\right\}\right)}{\text{Prob}\left(\phi \leq t^*\right)} \\
&= \frac{\text{Prob}\left(\left\{F(\phi) \geq p^* e^{-\lambda t}\right\} \cap \left\{F(\phi) \leq p^*\right\}\right)}{p^*} \\
&= \frac{\text{Prob}\left(F^{-1}(p^* e^{-\lambda t}) \leq \phi \leq F^{-1}(p^*)\right)}{p^*} \\
&= \frac{F\left(F^{-1}(p^*)\right) - F\left(F^{-1}(p^* e^{-\lambda t})\right)}{p^*} \\
&= \frac{p^*\left(1 - e^{-\lambda t}\right)}{p^*} \\
&= 1 - e^{-\lambda t},
\end{aligned}
$$

i.e. exponential with parameter λ. Moreover, because ψ is a strictly decreasing function of ϕ, the corresponding random variables are perfectly negatively correlated. As we now show this implies that the resulting mechanism is optimal for the multi-agent problem.

Theorem 3. *Under the mechanism described above, the probability that a bank run occurs on or before date t is given by*

$$
\text{Prob}(\max_i \tilde{r}_i \leq t) = \begin{cases} 0 & \textit{if } t \leq t^* + t^{**} \\ 1 - e^{-\lambda[t-(t^*+t^{**})]} & \textit{otherwise} \end{cases}
$$

where $t^{**} = \frac{1}{\lambda}\log(1 + p^*)$. Moreover this achieves the minimum among all feasible mechanisms.

The proof of the first part will be presented here, the proof of optimality can be found in Appendix B. Under the proposed mechanism, the run times for the two agents are deterministic functions of the state change date, ϕ. In particular,

$$\tilde{r}_1(\phi) = \phi + t^*$$
$$\tilde{r}_2(\phi) = \begin{cases} \phi & \text{if } \phi \geq t^* \\ -\frac{1}{\lambda}\log\left(\frac{F(\phi)}{p^*}\right) + t^* & \text{if } \phi < t^* \end{cases}$$

The graphs of these functions are depicted in Figure 6. Highlighted in blue is the graph of $\max \tilde{r}_i$, the date at which a bank run occurs.

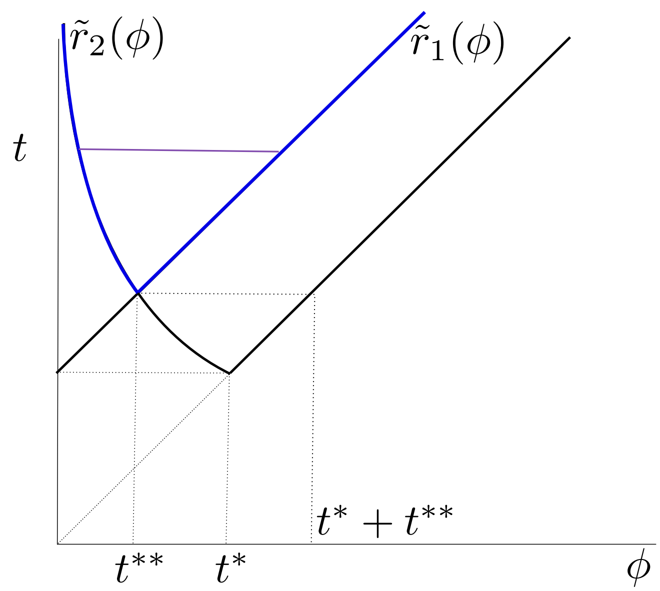


Figure 6: Beep times $\tilde{r}_i$ as a function of the state change time ϕ for the optimal multi-agent mechanism.

First, notice that a bank run cannot occur prior to the date $t^* + t^{**}$. The time t^{**} is the additional certain delay enjoyed by the principal from using private messages. We can calculate it from the fact that t^{**} is the unique value of ϕ at which the run times of the two agents are equal, i.e. $\tilde{r}_1(t^{**}) = \tilde{r}_2(t^{**})$. Since this occurs prior to t^* it is on the auxiliary clock segment of agent 2's mechanism and so

$$\begin{aligned}-\frac{1}{\lambda}\log\left(\frac{F(t^{**})}{p^*}\right)+t^* &= t^{**}+t^* \\ F(t^{**}) &= p^* e^{-\lambda t^{**}} \\ 1-e^{-\lambda t^{**}} &= p^* e^{-\lambda t^{**}} \\ t^{**} &= \frac{1}{\lambda}\log(1+p^*) \qquad (7)\end{aligned}$$

Next consider any date t and the event $\{\max \tilde{r}_i \leq t\}$. This event occurs whenever ϕ lands within the interval depicted in the figure, i.e.

$$\phi \in \left[\psi^{-1}(t), t-t^*\right]$$

The probability of this event is

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

Now from Equation 6,

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

hence

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

Since $F(t-t^*) = 1-e^{-\lambda(t-t^*)}$ we have

$$\begin{aligned}\text{Prob}(\max \tilde{r}_i \leq t) &= 1-e^{-\lambda(t-t^*)} - p^* e^{-\lambda(t-t^*)} \\ &= 1-e^{-\lambda(t-t^*)}(1+p^*)\end{aligned}$$

and by Equation 7, $1+p^* = e^{\lambda t^{**}}$, hence

$$\begin{aligned}&= 1-e^{-\lambda(t-t^*)}e^{\lambda t^{**}} \\ &= 1-e^{-\lambda[t-(t^*+t^{**})]}.\end{aligned}$$

The mechanism described above has the virtue of being deterministic but the drawback of being somewhat complicated, and in particular asymmetric. I conclude this section by describing a simpler symmetric mechanism that involves randomization. In this mechanism the principal secretly, randomly assigns the following two mechanisms to the agents. The first mechanism is beep-on, i.e. the agent who is assigned this mechanism will be informed immediately upon a state change. The second mechanism is a delay of length $t^* + t^{**}$. It can be shown that under this random mechanism, conditional on waiting each agent assigns a probability to `distressed` which increases and reaches the threshold p^* at date $t^* + t^{**}$ at which point he begins switching to `run` at rate λ. This achieves the optimum because with probability 1, the agent with the delayed mechanism always runs second and only after $t^* + t^{**}$.[27]

4 Long-Run/Strategic Agent

With some modifications the same machinery can be used to characterize the optimal mechanism when the (single) agent is a patient and strategic actor. Before beginning with the analysis, a few words about the effect on the principal of interacting with a strategic/long-run agent. First of all the principal benefits from the agent's patience in that it provides additional incentive leverage over the agent. The principal can induce the agent to take actions which are not short-run best-responses by promising more information in the future if the agent complies and threatening to punish with less information in the future if the agent refuses. On the other hand since varying the amount of information released in the future is the only instrument available to the principal, this incentive effect is of limited value. Indeed it is easy to characterize the incentive constraint. The most severe punishment with which the principal can threaten the agent is to keep the agent uninformed forever. Therefore, at any point in time the principal can induce the agent to take an action if and only if it gives the agent a continuation value at least as large as this autarky outcome.

We now consider a long-run payoff maximizing agent who shares the same discount factor as the principal. For simplicity in this section we

[27]It is interesting to observe that this mechanism is *not* individually optimal. One agent switches to `run` prior to t^* with positive probability. Thus, while the symmetric mechanism maximally forestalls the date at which both agents run, it would outperform the mechanism in the text if the principal had any more general payoff with losses experienced when a single agent runs.

will restrict attention to the two-state case, $S = \{0, 1\}$. The principal announces and commits to a mechanism and the agent chooses his actions to maximize his expected discounted average payoff. This yields an optimal strategy for the agent which is a choice of action after each possible history. Conversely a strategy for the agent is *incentive compatible* if there exists a mechanism such that after every possible history the agent maximizes his continuation payoff by following the strategy rather than deviating to an alternative strategy. In that case we can say that the mechanism *enforces* the strategy.

After any history, the agent's continuation value must be at least as large as the maximum value he can obtain when he remains uninformed for the remainder of the process. Moreover there always exists a mechanism under which the latter is the maximum continuation value available to the agent, namely the mechanism which never reveals any information. It follows that any incentive compatible strategy is enforceable by a mechanism that threatens to withhold all future information if the agent ever deviates from the strategy. Thus a strategy is enforceable if and only if there exists a mechanism under which the strategy achieves at least that minimum continuation value at every history. We will denote the minimum continuation value by $m(\mu)$, noting that it depends on the agent's current belief (and nothing else).

Characterizing the optimal mechanism for the principal will now require keeping track of pairs of continuation values (for the principal and agent.) To that end we modify the notation from before as follows. Let $\mathcal{U} : \Delta S \rightrightarrows \mathbf{R}^2$ represent the feasible flow value pairs given the current beliefs of the agent. In particular, $\mathcal{U}$ is the correspondence defined as follows.

$$\mathcal{U}(\mu) = \{(x, u) : \text{ There exists } a \text{ such that } x = x(a, \mu) \text{ and } u = u(a, \mu)\}$$

where we are writing, e.g. $x(a, \mu) = \mathbf{E}_\mu x(a, s)$. Note that the correspondence $\mathcal{U}$ has compact graph.

The maximum continuation value for the agent is achieved by a mechanism which always perfectly reveals the current state of the process. Let $M(\mu)$ denote this maximum continuation value when the agent begins with belief μ. The state space for the principal is now two-dimensional. The principal's value will depend on the current belief of the agent as well as the agent's promised continuation value, denoted w. Denote by $\mathcal{W} \subset [0, 1] \times \mathbf{R}$ the set of pairs (μ, w) such that $w \in [m(\mu), M(\mu)]$. A candidate value function for the principal is then a function

$$V : \mathcal{W} \to \mathbf{R}.$$

Associated with any such value function is a correspondence

$$V^{\mu} : [0,1] \rightrightarrows \mathbf{R}^2$$

which gives for each belief the set of value pairs (w, v) for agent and principal. In particular, $V^{\mu}(\mu) = \operatorname{graph} V(\mu, \cdot)$. Conversely, for any correspondence $g : [0,1] \rightrightarrows \mathbf{R}^2$, we can associate the correspondence $g^{\mathcal{W}} : \mathcal{W} \rightrightarrows \mathbf{R}$ given by $g^{\mathcal{W}}(\mu, w) = \{v : (w, v) \in g(\mu)\}$.

Consider the following operator on value functions V.

$$TV = \operatorname{cav} \left[(1-\delta)\mathcal{U} + \delta(V^{\mu} \circ f)\right]^{\mathcal{W}}.$$

Here the convex combination of correspondences $\mathcal{U}$ and $(V^{\mu} \circ f)$ is taken pointwise, so that its image at a point μ is the set $\{(1-\delta)(x, u) + \delta(w, v) : (x, u) \in \mathcal{U}(\mu) \text{ and } (w, v) \in V^{\mu}(f(\mu))\}$. The operation cav applied to a correspondence g is the pointwise smallest concave function on $\mathcal{W}$ which is pointwise no smaller than any element of the image of g.

Theorem 4. *1. The operator T is well defined. In particular TV exists for every function $V : \mathcal{W} \to \mathbf{R}$.*

2. The operator T is a contraction and has a unique fixed point V^*.

3. The value $V^*(\mu, w)$ gives the (continuation) value to the principal of the optimal incentive compatible mechanism when the agent has belief μ and receives continuation value w.

4. The optimal initial value for the principal when the agent begins with prior μ_0 is

$$\max_{\{w:(\mu_0, w) \in \mathcal{W}\}} V^*(\mu_0, w).$$

4.1 An Example

Consider the following dynamic principal-agent problem. An agent decides in each period whether to work or shirk. His payoff is zero in any period that he works. His payoff from shirking depends on the state of the world. The principal always prefers that the agent work. The principal's payoff is 1 in any period that the agent works and 0 in any period that the agent shirks.

The state of the world is binary $s \in \{0,1\}$ and indicates whether the firm is in decline with $s = 1$ indicating that the firm will soon shut down. The more likely it is that the firm will shut down the less willing the agent is to work as his career prospects within the firm are disappearing. To model this, assume that the agent's payoff from shirking is -1 in state $s = 0$ and 1 in state $s = 1$. In particular, the agent prefers to shirk whenever the probability of state $s = 1$ is at least $1/2$.

The firm observes the state and can send status reports to the agent. It is common knowledge that the firm begins in state $s = 0$ and transitions to state 1 with probability $\Lambda = (1 - e^{-\lambda\Delta})$ where λ is a continuous-time transition parameter and Δ is the real-time length of a period. State $s = 1$ is absorbing. Let $f(\mu) = \mu + (1 - \mu)\Lambda$. The principal and agent discount the future at equal discount rates $r > 0$, yielding discrete-time discount factor $\delta = e^{-r\Delta}$.

Note that if the agent were myopic and always maximized the current payoff, his decision rule would be to work whenever $\mu \leq 1/2$ and shirk otherwise. Thus the problem would be identical to the email beeps problem and the principal's optimal mechanism would be to use a delayed beep as characterized previously. Hereafter I will refer to that mechanism as the *static-optimum.* I will show that even when the agent is patient the static optimal mechanism remains optimal. (It will no longer be uniquely optimal however.)

The key observation is the following.

Lemma 4. *The static optimal mechanism minmaxes the agent. In particular, at every belief μ, the agent's continuation value when the principal uses the static optimal mechanism is the same as when the principal never gives any information about the state, namely $m(\mu)$.*

Note that the principal's value from the static-optimum is strictly greater than what he would earn if he were to keep the agent uninformed. Thus the lemma implies that the static optimum represents one way to increase the principal's payoff while keeping the agent at his individually rational level. The remaining question is then whether, using a more complicated mechanism the principal can do even better than this without violating the individual rationality constraint. In particular when the agent is patient the principal could try to induce him to work even at beliefs greater than $\mu = 1/2$ by offering the promise of better information in the future as a reward. I now show that this is impossible, any dynamic incentive scheme costs the principal (weakly) more in the long run than he gains from prolonging the agent's effort beyond $\mu = 1/2$.

Proposition 2. *For every $\mu \in [0,1]$, let $\underline{v}(\mu)$ and $\bar{v}(\mu)$ be the principal's values from the full-information and static-optimal mechanism respectively. For each $(\mu, w) \in \mathcal{W}$, w is a convex combination of $m(\mu)$ and $M(\mu)$ with weights α and $(1-\alpha)$, and the principal's value function for the optimal dynamic mechanism is linear:*

$$V^*(\mu, w) = \alpha \bar{v}(\mu) + (1-\alpha)\underline{v}(\mu)$$

Since at every belief $\mu \in [0,1]$ this is maximized by keeping the agent at his minmax value, it follows that the static optimum is an optimal mechanism when the agent is patient and strategic.

The proof is in Section D.2.

5 Continuous Time Analysis

Up to this point I have analyzed each example by first considering a discrete-time optimization and then taking continuous-time limits. In many cases it is more convenient to conduct the analysis directly in continuous time. Let r denote the continuous time discount rate and recall that $\dot{\mu}$ is the continuous-time law of motion for the agent's beliefs absent any further information for the principal. In the appendix I derive the Hamilton-Jacobi-Bellman (HJB) equation which we can express as a functional equation as follows:

$$rV = \text{cav}\left[u + V' \cdot \dot{\mu}\right].$$

Like in the discrete-time version the concavification operator facilitates a useful geometric representation, but now the optimal value function is expressed in terms of its first derivative and the continuous time law of motion.

In this section I will demonstrate the usefulness of the continuous-time formulation by solving the 3-action email problem through a series of diagrams. For convenience let's normalize the discount rate r to 1. To begin with, let's verify that the continuous-time limit value function from the email beep problem verifies the HJB equation. Recall that the optimal policy is to wait until the agent's belief reaches p^* before sending messages and after that to send messages that randomize the agent's beliefs between p^* and 1. This yields the following value function in continuous time

$$V^*(\mu) = \begin{cases} \int_0^{t(\mu)} e^{-t}dt + e^{-t(\mu)}V^*(p^*) & \text{for } \mu \leq p^* \\ (1-\mu)V^*(p^*) & \end{cases}$$

where $t(\mu)$ is the time required for beliefs to evolve from μ to p^*,

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

and $V(p^*)$ is the continuation value at the threshold which we previously calculated to be $1/(1+\lambda) < 1$.

To calculate the right-hand-side of the HJB equation we need to compute $V' \cdot \dot{\mu}$. The simplest way to do this is via a change of variables expressing the value function in terms of time rather than beliefs. Let $t^* = t(0)$ be the total time it takes for the agent's beliefs to first reach p^*, and for all t^*:

$$W(t) = \int_0^{t^*-t} e^{-s}ds + e^{-(t^*-t)}W(t^*)$$

Then $V' \cdot \dot{\mu}$ is just the derivative with respect to t:

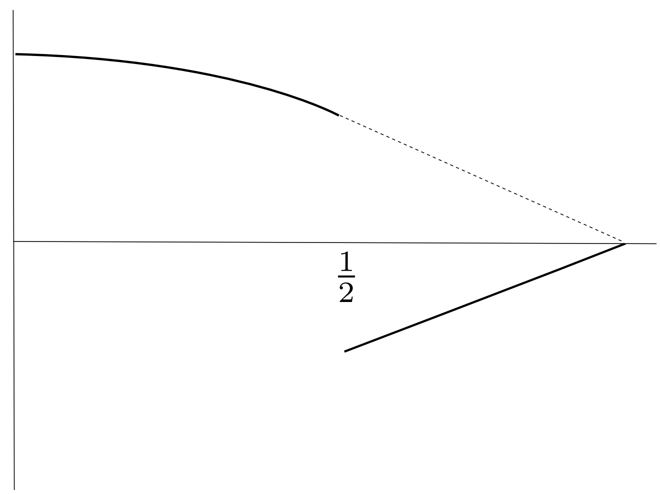
$$\frac{dW}{dt}(t) = e^{-(t^*-t)}\left[W(t^*) - 1\right]$$

(As time passes, the continuation value falls because the point in time draws closer when the principal will stop earning flow payoff 1 and instead earn continuation value $W(t^*) = V(p^*) < 1$.)

Since $t^* - t = t(\mu)$, to the left of p^*, the expression $u + V' \cdot \dot{\mu} = u + \frac{dW}{dt}(t)$ is given by

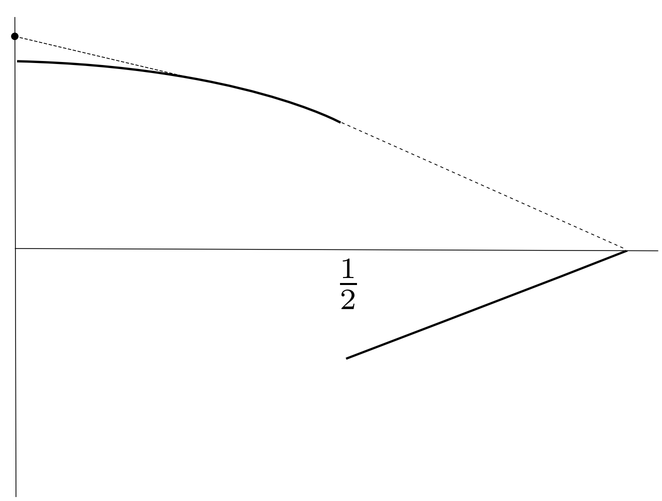
$$(1 - e^{-t(\mu)}) + e^{-t(\mu)}V(p^*)$$

and to the right, since the value function is linearly decreasing and the flow payoff is zero, $V' \cdot \dot{\mu}$ is proportional to $-\dot{\mu} = -\lambda(1-\mu)$. Thus, the bracketed expression on the right-hand side of the HJB equation has the following shape,



and when we concavify we indeed recover the value function as the HJB equation requires. Moreover we see that the geometry of the HJB equation reveals the optimal mechanism in the same way as we saw for the concavification of the discrete-time Bellman equation.

With these observations in hand we can now show formally how the optimal mechanism changes in the three action version. Indeed, if we take V as a candidate continuous-time value function for the three-action problem, the HJB equation tells us to compute its time-derivative and add it to the three-step u in Equation 4. We obtain exactly the same graph except that the height at $\mu = 0$ jumps up by $1/4$. When we concavify:



We see that we do not recover V but instead there is now a linear segment over an initial interval. This gives us a hint that the optimal value function will have a similar shape. And indeed we can reduce the functional fixed-point problem in the HJB equation to a parametric equation with a single unknown p^{**}. For suppose we pick a threshold p^{**} where the mechanism switches from randomizing to silence. Knowing that the value function will be smooth at that point tells us what the slope of the initial linear segment must be. It must be the slope of the email value function at p^{**}, namely $V'(p^{**})$. This therefore tells us a candidate value for $V(0)$:

$$V(p^{**}) + p^{**}V'(p^{**})$$

Now when we take the resulting candidate value function and feed it into the right-hand side of the HJB equation we obtain a new value for $V(0)$:

$$u(0) + V'(p^{**})\lambda$$

We can solve the model by picking the p^{**} that equates these and thus produces a fixed point.

$$V(p^{**}) + p^{**}V'(p^{**}) = 5/4 + V'(p^{**})\lambda.$$

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.

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.

Bimpikis, Kostas, and Kimon Drakopoulos. 2014. "Disclosing Information in Strategic Experimentation."

Brandenburger, Adam, and Eddie Dekel. 1993. "Hierarchies of beliefs and common knowledge." *Journal of Economic Theory*, 59(1): 189–198.

Cardaliaguet, Pierre, Catherine Rainer, Dinah Rosenberg, and Nicolas Vieille. 2013. "Markov games with frequent actions and incomplete information." *arXiv preprint arXiv:1307.3365*.

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, and Marcin Peski. 2004. "Hierarchies of belief and interim rationalizability." *CMS-EMS Northwestern University Working Paper*, , (1388).

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

Guo, Yingni, and Johannes Hörner. 2014. "Dynamic Mechanissms Without Money." Yale University.

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

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

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

Mertens, Jean-Francois, and Shmuel Zamir. 1985. "Formulation of Bayesian analysis for games with incomplete information." *International Journal of Game Theory*, 14(1): 1–29.

Morris, Stephen, and Hyun Song Shin. 1999. "7 A theory of the onset of currency attacks." *The Asian Financial Crisis: Causes, Contagion and Consequences*, 2: 230.

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*.

Renault, Jerome, Eilon Solan, and Nicolas Vielle. 2014. "Optimal Dynamic Information Provision."

Rockafellar, R Tyrrell. 1997. *Convex analysis.* Princeton university press.

Sorin, Sylvain. 2002. *A first course on zero-sum repeated games.* Vol. 37, Springer.

A Proofs for Section 2

Proof of the Obfuscation Principle. T hat a policy induces a stochastic process satisfying the three 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, the second condition states that the agent's Bayesian belief ν_t is the correct conditional probability distribution over states, and the third condition reflects the agent's knowledge of the underlying stochastic process governing the state.

Given item 2, the principal's expected payoff conditional on inducing interim belief ν_t is just $\sum_{s \in S} \nu_t(s) u(\nu_t, s)$, i.e. $u(\nu_t)$ justifying the expression for the principal's long-run expected payoff in the statement of the lemma.

To prove the converse, let (μ_t, s_t, ν_t) be any stochastic process satisfying item 1, item 2 and item 3. 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 $B \subset \Delta(M)$,

$$\int_{\nu \in B} \nu(s) dZ \leq \int_{\nu \in \Delta(M)} \nu(s) dZ = \mathbf{E}\nu(s)$$

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 $B \in \Delta(S)$,

$$\int_{\nu \in B} y(\nu) dZ(\nu) = \mu_t(s)\sigma_s(B) = \text{Prob}(\{s\} \times B)$$

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. item 2 follows as well. □

Proof of Theorem 1. Consider the operator

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

on the complete metric space of bounded real-valued functions with the sup metric. The Blackwell sufficient conditions for T to be contraction mapping (with modulus δ) are as follows.

1. If $V \geq V'$ then $TV \geq TV'$,
2. If c is a constant then $T(V+c) \leq TV + \delta c$.

The first follows from the observation that if $g \geq g'$, then any concave function which is weakly larger than g is also weakly larger than g' and therefore $\text{cav}\, g \geq \text{cav}\, g'$. The second follows from the observation that for any function h and any constant k, a concave function y is weakly larger than h if and only if $y+k$ is weakly larger than $h+k$. Therefore $\text{cav}(h+k) = (\text{cav}\, h) + k$. □

Proof of Theorem 2. The value function is as follows.

$$V(\mu) = (1-\delta)\left[\sum_{s=0}^{n(\mu)-1} \delta^s + \delta^{n(\mu)}\left(\frac{1-f^{n(\mu)}(\mu)}{1-p^*}\left(1 + \frac{\delta(1-p^{**})}{1-p^*-\delta(1-p^{**})}\right)\right)\right]$$

where

$$n(\mu) = \min\{n \geq 0 : f^n(\mu) > p^*\}.$$

We will now prove that V so defined is the value associated with the mechanism described above and that it satisfies the optimality condition in Equation 2. Let

$$Z(\mu) = (1-\delta)u(\mu) + \delta V(f(\mu)).$$

According to Equation 2, to show that V is the optimal value function we must show that V is the concavification of Z. The latter is defined as the concave function which is the pointwise minimum of all concave functions that are pointwise no smaller than Z.

First note that V as defined is a concave function. It consists of $n(0)$ linear segments with decreasing slope to the left of p^*, followed by a linear segment from p^* to 1. See Figure 1. We next show that for $\mu \leq p^*$,

$$V(\mu) = Z(\mu).$$

To do this, write $\mu^+ = f(\mu)$. Since $u(\mu) = 1$ when $\mu \leq p^*$,

$$
\begin{aligned}
&(1-\delta)u(\mu) + \delta V(\mu^+) \\
&\quad = (1-\delta) + \delta\,(1-\delta)\left[\sum_{s=0}^{n(\mu^+)-1}\delta^s + \delta^{n(\mu^+)}\left(\frac{1-f^{n(\mu^+)}(\mu^+)}{1-p^*}\left(1+\frac{\delta(1-p^{**})}{1-p^*-\delta(1-p^{**})}\right)\right)\right] \\
&\quad = (1-\delta)\left\{1+\delta\left[\sum_{s=0}^{n(\mu^+)-1}\delta^s + \delta^{n(\mu^+)}\left(\frac{1-f^{n(\mu^+)}(\mu^+)}{1-p^*}\left(1+\frac{\delta(1-p^{**})}{1-p^*-\delta(1-p^{**})}\right)\right)\right]\right\} \\
&\quad = (1-\delta)\left\{\sum_{s=0}^{n(\mu^+)}\delta^s + \delta^{n(\mu^+)+1}\left(\frac{1-f^{n(\mu^+)+1}\mu}{1-p^*}\left(1+\frac{\delta(1-p^{**})}{1-p^*-\delta(1-p^{**})}\right)\right)\right\} \\
&\quad = V(\mu)
\end{aligned}
$$

where the last equality follows because $n(\mu) = n(\mu^+) + 1$.

So far we have shown that V is concave and coincides with Z for all $\mu \leq p^*$. Since the optimal value function is the convex hull of Z which is defined as the pointwise minimum of all concave functions that are no smaller than Z, we have shown that $V(\mu)$ is the optimal value for all $\mu \leq p^*$.

Now consider $\mu \in (p^*, 1]$. First observe that $V(1) = 0 = Z(1)$. We will now show that any concave function which is pointwise no smaller than Z must also be pointwise no smaller than V for all $\mu \in (p^*, 1]$. Let Y be any concave function such that $Y(p^*) \geq Z(p^*)$ and $Y(1) \geq Z(1)$. Then by concavity, for $\mu = \alpha p^* + (1-\alpha)1$ for $\alpha \in (0, 1)$,

$$Y(\mu) \geq \alpha Z(p^*) + (1-\alpha)Z(1) = \alpha V(p^*) + (1-\alpha)V(1)$$

and the latter is exactly $V(\mu)$ since V is linear over $\mu \in [p^*, 1]$.

This concludes the proof that V is the optimal value function. Next we show that V is the value function for the specified mechanism. First consider the belief p^{**}. The mechanism specifies that the principal randomizes over two interim beliefs, p^* and 1. Let α be the probability of interim belief p^*. By the martingale property

$$\alpha p^* + (1-\alpha) = p^{**}$$

so that

$$\alpha = \frac{1-p^{**}}{1-p^*}.$$

Thus, if the mechanism generates value function W, then

$$W(p^{**}) = \alpha \left[(1-\delta) + \delta W(f(p^*))\right]$$

because with probability α, the principal staves off email checking for one period after which the belief is updated to $f(p^*)$ (and with the remaining probability the agent checks email and the game ends.) Since $p^{**} = f(p^*)$,

$$W(p^{**}) = \frac{(1-\delta)(1-p^{**})}{1-p^*}\left(1 + \frac{\delta}{1-\delta}W(p^{**})\right)$$

implying that

$$W(p^{**}) = \frac{(1-\delta)(1-p^{**})}{1-p^* - \delta(1-p^{**})} \tag{9}$$

and plugging p^{**} into V verifies that $V(p^{**})$ equals the right-hand side.

Moreover, for any $\mu > p^*$, the mechanism yields value

$$\frac{(1-\delta)(1-\mu)}{1-p^*}\left(1 + \frac{\delta}{1-\delta}V(p^{**})\right)$$

which is $V(\mu)$. Finally if $\mu \leq p^*$ the mechanism delays checking for $n(\mu)$ periods after which a belief above p^* is reached and hence the value is $V(\mu)$.

Continuous Time Limit Consider the continuous time limit as $\Delta \to 0$. First, note that the discrete time transition probability, f from state 0 to state 1 is $1 - e^{-\lambda\Delta}$. Consider the belief p^{**}. Rewrite Equation 9 as follows

$$V(p^{**}) = \frac{(1-\delta)\frac{1-p^{**}}{1-p^*}}{1-\delta\frac{1-p^{**}}{1-p^*}}.$$

Since

$$p^{**} = p^* + (1-p^*)(1-e^{-\lambda\Delta})$$

we have

$$1 - p^{**} = (1-p^*)e^{-\lambda\Delta}$$

so

$$\frac{1-p^{**}}{1-p^*} = e^{-\lambda\Delta}$$

which yields

$$V(p^{**}) = \frac{(1-e^{-r\Delta})e^{-\lambda\Delta}}{1-e^{-r\Delta}e^{-\lambda\Delta}}$$

and applying l'Hopital's rule ,

$$\begin{aligned}\lim_{\Delta\to 0} V(p^{**}) &= \lim_{\Delta\to 0} \frac{e^{-\lambda\Delta}-e^{-(\lambda+r)\Delta}}{1-e^{-(\lambda+r)\Delta}}\\ &= \frac{-\lambda+r+\lambda}{r+\lambda}\\ &= \frac{r}{r+\lambda}\end{aligned}$$

Note that $r/(r+\lambda)$ is equal to the average discounted value of a mechanism which immediately notifies the user when an email arrives. To see this, note that the latter value is equal to the expected discounted waiting time before the first email arrival, i.e.

$$\begin{aligned}\int_0^\infty \int_0^t re^{-rs}ds\lambda e^{-\lambda t}dt &= \int_0^\infty \left[e^{-rs}|_{s=0}^{s=t}\right]\lambda e^{-\lambda t}dt\\ &= \int_0^\infty (1-e^{-rt})\lambda e^{-\lambda t}dt\\ &= 1-\int_0^\infty e^{-rt}\lambda e^{-\lambda t}dt\\ &= 1-\lambda\int_0^\infty e^{-(r+\lambda)t}dt\\ &= 1-\left[\frac{\lambda}{-(r+\lambda)}e^{-(r+\lambda)t}|_0^\infty\right]\\ &= 1+0-\frac{\lambda}{r+\lambda}\\ &= \frac{r}{r+\lambda}\end{aligned}$$

Now $p^{**} \to p^*$ as $\Delta \to 0$. Thus, in the limit $V(p^*)$ equals the *initial* value of a mechanism that notifies the user immediately when an email arrives. It follows that $V(p^*)$ can also be attained by a mechanism which only notifies the user whether an email had arrived t^* units of time *in the past* where

$$1-e^{-\lambda t^*} = p^*,$$

i.e. exactly the amount of time it takes for the cumulative probability of an email arrival to reach the threshold p^*. Indeed we will now show that the limiting value function at all beliefs below p^* is identical to the value function of such a mechanism.

When the principal delays notification for t^* units of time, no information is revealed to the agent until time t^*, at which point the agent's belief is p^*. When the agent's belief at some time t equals p^* and there is no email beep the agent learns that no email arrived prior to time $t - t^*$ and obtains no information about any arrival in the most recent t^*-length interval of time. By the definition of t^*, the probability of an arrival during that period is exactly p^*. Thus, in the absence of any beep, the agent's belief remains constant at p^*.

This means that a belief $\mu < p^*$ occurs exactly once; namely at the time $\tau < t^*$ such that

$$1 - e^{-\lambda\tau} = \mu.$$

Beginning at time τ, the principal's payoff from the delayed-beep mechanism is 1 until time t^*, after which he obtains continuation value $V(p^*)$. This yields discounted expected value

$$\left(1 - e^{-r(t^*-\tau)}\right) + e^{-r(t^*-\tau)}V(p^*).$$

Turning now to the optimal value function, since

$$\lim_{\Delta\to 0} f^{n(\mu)}(\mu) = p^*,$$

the continuous-time limit value beginning at τ is also

$$\lim_{\Delta\to 0} V(\mu) = \left(1 - e^{-r(t^*-\tau)}\right) + e^{-r(t^*-\tau)}V(p^*).$$

and thus the delayed-beep mechanism is optimal. □

B Proofs for Section 3

Proof of Lemma 3. To begin with, notice that

$$\rho^1_{-i} = \{\beta_{-i} : \beta_{-i}(\{\texttt{healthy}\} \times H_i) < 1 - p^*\} \subset \left\{\beta_{-i} : \beta_{-i}(\{\texttt{healthy}\} \times \neg\rho^1_i) < 1 - p^*\right\} = \rho^2_{-i}$$

since of course $\neg\rho_i^1 \subset H_i$. Now, for the purposes of induction, suppose that $\rho_i^k \subset \rho_i^{k+1}$. Then

$$\rho_{-i}^{k+1} = \left\{\beta_{-i} : \beta_{-i}(\{\texttt{healthy}\} \times \neg\rho_i^k) < 1 - p^*\right\}$$
$$\subset \left\{\beta_{-i} : \beta_{-i}(\{\texttt{healthy}\} \times \neg\rho_i^{k+1}) < 1 - p^*\right\} = \rho_{-i}^{k+2}$$

and we have shown that the sequence of sets $\{\rho_i^k\}$ is nested (expanding).

Consider the probability

$$\beta_i(\{\texttt{healthy}\} \times \neg\rho_{-i}) = \beta_i(\{\texttt{healthy}\} \times \neg \cup_{k\geq 1} \rho_{-i}^k)$$
$$= \beta_i(\{\texttt{healthy}\} \times \cap_{k\geq 1}\neg\rho_{-i}^k)$$

We have

$$\neg\rho_{-i}^k \subset \neg\rho_{-i}^{k-1}$$

and hence

$$\{\texttt{healthy}\} \times \neg\rho_{-i}^k \subset \{\texttt{healthy}\} \times \neg\rho_{-i}^{k-1}$$

hence $\beta_i(\{\texttt{healthy}\} \times \neg\rho_{-i}) < 1 - p^*$ if and only if for some k we have $\beta_i(\{\texttt{healthy}\} \times \neg\rho_{-i}^k)$, i.e. if and only if $\beta_i \in \rho_i^k$ for some k, i.e. if and only if $\beta_i \in \cup_k\rho_i^k = \rho_i$. □

Proof of Theorem 3. 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`. We will show that the mechanism described in the text achieves the maximum value for the relaxed problem. Because the necessary conditions for an agent to wait are weaker in the relaxed problem the value of the relaxed problem is no larger than the value of the original problem this will prove the theorem.

In the relaxed problem there is no loss of generality in restricting attention to individually optimal mechanisms. Increasing the time before agent i runs can only increase $\max \tilde{r}_i$ and (in the relaxed problem) cannot affect the incentives of the other agent. We have shown that an individually optimal mechanism is a "beep" mechanism: the agent either hears a beep or silence, a beep signals that the bank is distressed with probability 1, and an agent runs if and only if he hears a beep. 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,

$$\text{Prob}(\phi \leq t \mid \tilde{r}_i > t) \leq p^*. \tag{10}$$

We can express the conditional probability on the left-hand side as follows.[28]

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

and thus we can re-arrange the inequality in Equation 10

$$1 + \frac{\text{Prob}(\phi > t)}{\text{Prob}(\phi \leq t) - \text{Prob}(\tilde{r}_i \leq t)} \geq \frac{1}{p^*}$$

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

$$\text{Prob}(\tilde{r}_i \leq t) = \begin{cases} 0 & \text{if } t \leq t^* \\ \left(\frac{1}{1-p^*}\right) [\text{Prob}(\phi \leq t) - p^*] & \text{otherwise.} \end{cases} \tag{11}$$

Next consider

$$\text{Prob}(\tilde{r}_2 \leq t \mid \tilde{r}_1 > 1),$$

[28]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. Similar reasoning explains the denominator in the second line.

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[29]

$$\text{Prob}(\tilde{r}_2 \leq t \mid \tilde{r}_1 > t) \leq \text{Prob}(\phi \leq t \mid \tilde{r}_1 > t),$$

and since by Equation 10, 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 11) and thus

$$\text{Prob}(\max \tilde{r}_i \leq t) \geq (1 + p^*)\text{Prob}(\tilde{r}_2 \leq t) - p^*.$$

Using Equation 11 again we obtain

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

or

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

Recall that

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

[29]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.

allowing the previous inequality to be rewritten

$$\text{Prob}(\max \tilde{r}_i \leq t) \geq e^{\lambda(t^*+t^{**})}\text{Prob}(\phi \leq t) + (1 - e^{\lambda(t^*+t^{**})})$$

and since ϕ is distributed exponentially with parameter λ,

$$\begin{aligned}\text{Prob}(\max \tilde{r}_i \leq t) &\geq e^{\lambda(t^*+t^{**})}\left(1 - e^{-\lambda t}\right) + (1 - e^{\lambda(t^*+t^{**})}) \\ &= 1 - e^{-\lambda(t-[t^*+t^{**}])}.\end{aligned}$$

□

C Deriving The HJB Equation

Let us approximate the optimized continuous time discounted payoff for the principal by discretizing the time dimension into Δ intervals and the summation

$$J(t, \mu_t) = \mathbf{E}\sum_{s=0}^{\infty} e^{-r(t+s\Delta)}u(\mu_{t+s\Delta}) \cdot \Delta + O(\Delta^2).$$

Here $J(t, \mu_t)$ gives the principal's maximal expected total discounted continuation payoff beginning at time t when the agent's beliefs at instants $\{t + s\Delta\}_{s\geq 0}$ are given by $\mu_{t+s\Delta}$. By the principal of optimality

$$J(t, \mu_t) = \max_{\substack{p\in\Delta(\Delta S)\\ \mathbf{E}p=\mu_t}} \left[e^{-rt}u(\nu_t)\Delta + e^{-rt}J(t + \Delta, f(\nu_t))\right] + O(\Delta^2).$$

where p denotes a lottery whose realization is ν_t. The optimal policy is stationary so $J(t, \mu) = J(t', \mu)$ and we can write

$$J(t, \mu) = e^{-rt}V(\mu)$$

and

$$e^{-rt}V(\mu_t) = \max_{\substack{p\in\Delta(\Delta S)\\ \mathbf{E}p=\mu_t}} \mathbf{E}_p\left[e^{-rt}u(\nu_t)\Delta + e^{-r(t+\Delta)}V(f(\nu_t))\right] + O(\Delta^2).$$

Dividing through by e^{-rt},

$$V(\mu_t) = \max_{\substack{p\in\Delta(\Delta S)\\ \mathbf{E}p=\mu_t}} \mathbf{E}_p\left[u(\nu_t)\Delta + e^{-r\Delta}V(f(\nu_t))\right] + O(\Delta^2).$$

Now a first-order approximation

$$e^{-r\Delta}V(f(\nu_t)) = e^{-r\Delta}\left[V(\nu_t) + V'(\nu_t)\dot{\nu}_t\Delta\right] + O(\Delta^2),$$

so

$$V(\mu_t) = \max_{\substack{p\in\Delta(\Delta S)\\ \mathbf{E}p=\mu_t}} \mathbf{E}_p\left\{u(\nu_t)\Delta + e^{-r\Delta}\left[V(\nu_t) + V'(\nu_t)\dot{\nu}_t\Delta\right]\right\} + O(\Delta^2).$$

I claim that at an optimum p^* of the maximization above, $\mathbf{E}_{p^*}V(\nu_t) = V(\mu_t)$. To see why, for each $\nu \in \Delta(S)$, let $p^*(\nu)$ be a maximizer for the optimization that defines $V(\nu)$. Then we have

$$u(\nu)\Delta + e^{-r\Delta}V(f(\nu)) \leq \mathbf{E}_{p^*(\nu)}\left[u(\nu')\Delta + e^{-r\Delta}V(f(\nu'))\right] = \max_{\substack{p\in\Delta(\Delta S)\\ \mathbf{E}p=\nu}} \mathbf{E}_p\left[u(\nu')\Delta + e^{-r\Delta}V(f(\nu'))\right]$$

since the left-hand side is a feasible value for the right-hand side optimization, taking p to be the degenerate lottery. Therefore

$$V(\mu) = \mathbf{E}_{p^*}\left[u(\nu)\Delta + e^{-r\Delta}V(f(\nu))\right] \leq \mathbf{E}_{p^*}\mathbf{E}_{p^*(\nu)}\left[u(\nu')\Delta + e^{-r\Delta}V(f(\nu'))\right] = \mathbf{E}_{p^*}V(\nu)$$

but also

$$V(\mu) \geq \mathbf{E}_{p^*}V(\nu)$$

since the compound lottery in the middle expression above is feasible for the optimization that defines $V(\mu)$. We can thus re-arrange as follows

$$\left(1 - e^{-r\Delta}\right)V(\mu_t) = \max_{\substack{p\in\Delta(\Delta S)\\ \mathbf{E}p=\mu_t}} \mathbf{E}_p\left[u(\nu_t)\Delta + e^{-r\Delta}V'(\nu_t)\dot{\nu}_t\Delta\right] + O(\Delta^2).$$

Dividing through by Δ and then taking $\Delta \to 0$ we obtain by l'Hopital's rule ,

$$rV(\mu_t) = \max_{\substack{p\in\Delta(\Delta S)\\ \mathbf{E}p=\mu_t}} \mathbf{E}_p\left[u(\nu_t) + V'(\mu_t)\dot{\mu}_t\right]$$

or

$$rV = \text{cav}\left[u + V'\dot{\mu}\right].$$

D Proofs for Section 4

Proof of Theorem 4. To prove item 1, define

$$Z = \left[(1-\delta)\mathcal{U} + \delta(V^{\mu} \circ f)\right]^{\mathcal{W}}$$

We will establish the following.

1. $Z(0, M(0)) \neq \emptyset$
2. $Z(1, M(1)) \neq \emptyset$
3. $Z(\mu, m(\mu)) \neq \emptyset$ for all μ.

To prove the first two, recall that $M(\cdot)$ is the agent's value from a mechanism that keeps him perfectly informed. When the agent's beliefs are degenerate ($\mu = 0$ or $\mu = 1$), there is no information to give to the agent and therefore, e.g.

$$M(0) = (1-\delta)\max_{a} w(a, 0) + \delta M(f(0))$$

the agent's value is his flow value from the optimal action today plus the discounted continuation value associated with the full-information mechanism. Therefore since

$$(M(f(0)), V(f(0), M(f(0)))) \in V^{\mu}(f(\mu))$$

and

$$(w(a^*, 0), u(a^*, 0)) \in \mathcal{U}(0)$$

where $a^* \in \operatorname{argmax}_a w(a, 0)$, we have that the convex combination of these points belongs to

$$(1-\delta)\mathcal{U}(0) + \delta(V \circ f)(0)$$

and that convex combination is

$$(M(0), v)$$

for $v = (1-\delta)u(a^*, 0) + \delta V(f(0), M(f(0)))$. In particular, $Z(0, M(0)) \neq \emptyset$. The identical argument establishes the second point.

For the third point, recall that $m(\mu)$ is the value to the agent from a mechanism which never reveals any information. By the principal of optimality,

$$m(\mu) = (1-\delta)\max_a x(a,\mu) + \delta m(f(\mu)).$$

The remainder follows the same lines as above.

Now define $\operatorname{co} Z : \mathcal{W} \rightrightarrows \mathbf{R}$ to be the correspondence whose graph is the convex hull of the graph of Z. That is,

$$\operatorname{graph} \operatorname{co} Z = \operatorname{conhull} \operatorname{graph} Z.$$

We will show that $\operatorname{co} Z$ has non-empty values, $\operatorname{co} Z(\mu, w) \neq \emptyset$ for all $(\mu, w) \in \mathcal{W}$. By the three points proven above it is enough to show that

$$\mathcal{W} = \operatorname{conhull}\left[\{(0, M(0)), (1, M(1))\} \cup \{(\mu, m(\mu)) : \mu \in [0,1]\}\right].$$

This follows from the fact that for any $\mu \in [0,1]$,

$$M(\mu) = \mu M(1) + (1-\mu) M(0),$$

i.e. the agent's full information value is achieved by sending a message which with probability μ gives him the full information value associated with belief 1 and with the remaining probability gives him the full information value associated with belief 0.

Finally, we can conclude the proof of item 1. For a given correspondence g let S_g be the set of all concave functions that are no smaller than g. That is, $h \in S_g$ if and only if h is concave and $h(\mu, w) \leq v$ for all $v \in g(\mu, w)$. By construction $Z(\mu, w) \subset \operatorname{co} Z(\mu, w)$ for all (μ, w) therefore $S_{\operatorname{co} Z} \subset S_Z$. By definition, $TV = \operatorname{cav} Z$ is the minimal element of S_Z. If we can show that TV also belongs to $S_{\operatorname{co} Z}$ we will have shown that TV is also the minimal element of $S_{\operatorname{co} Z}$. The minimal element of $S_{\operatorname{co} Z}$ clearly exists because we have shown that $\operatorname{co} Z$ has non-empty values bounded above by the maximum feasible flow payoff for the principal.[30]

That $TV \in S_{\operatorname{co} Z}$ follows immediately from concavity. Let $v = \operatorname{co} Z(\mu, w)$. Then (μ, w, v) belongs to the convex hull of the graph of Z. Then there exist points $z' = (\mu', w', v')$ and $z'' = (\mu'', w'', v'')$ belonging to the graph of Z

[30]The set of functions in $S_{\operatorname{co} Z}$ is therefore not empty. Since the pointwise infimum of concave functions is concave, the pointwise infimum of functions in $S_{\operatorname{co} Z}$ is an element of $S_{\operatorname{co} Z}$ and indeed the minimal element.

such that $(\mu, w, v) = (1-\alpha)z' + \alpha z''$. Since TV belongs to S_Z, $(TV)(\mu', w') \geq v'$ and $(TV)(\mu'', w'') \geq v''$. By concavity of TV, we then have $TV(\mu, w) \geq (1-\alpha)v' + \alpha v'' = v$. We have shown that $TV(\mu, w) \geq \operatorname{co} Z(\mu, w)$ for all (μ, w), i.e. that $TV \in S_{\operatorname{co} Z}$.

Turning to item 2, note that v belongs to $Z(\mu, w)$ if and only if there exists $(x, u) \in \mathcal{U}(\mu)$ and $w' \in [m(f(\mu)), M(f(\mu))]$ such that for $v' = V(f(\mu), w')$,

$$(w, v) = (1-\delta)(x, u) + \delta(w', v').$$

From this it follows that if $V \geq V'$, then $\max Z(\mu, w) \geq \max Z'(\mu, w)$ for all $(\mu, w) \in \mathcal{W}$. Moreover if c is a constant and Z^c is the correspondence derived from $V + c$ then $\max Z^c(\mu, w) = \max Z(\mu, w) + \delta c$. These observations combined with Theorem 1 imply item 2.

Next item 3. Begin with some definitions and notation. For correspondences $W : [0,1] \rightrightarrows \mathbf{R}^2$, define the operator B as follows

$$\begin{aligned} B(W)(\nu) = \{(w, v) : (w, v) = (1-\delta)(x, u) + \delta(w', v') \\ \text{for } (x, u) \in \mathcal{U}(\nu), (w', v') \in W(f(\nu)); w \geq m(\nu)\} \end{aligned}$$

and the operator C

$$\begin{aligned} C(W)(\mu) = \{(w, v) : \text{ There exists a lottery } (\tilde{\nu}, \tilde{w}, \tilde{v}) \text{ on graph } W \\ \text{such that } \mathbf{E}(\tilde{\nu}, \tilde{w}, \tilde{v}) = (\mu, w, v)\} \end{aligned}$$

and say that a correspondence W is self-generating if it is a fixed-point of $C \circ B$, i.e. $W \subset C(B(W))$.

Fix a mechanism σ and an optimal strategy α for the agent. An *interim history* is a history of play at which the principal has just sent a message but the agent has yet to take an action. A prior history is a history of play after which the agent has updated his interim belief ν to $f(\nu)$ but before the principal has sent a new message. Say that (w, v) is an *interim value pair* if there is an interim history such that payoff pair starting at that history is (w, v). Say that (w, v) is a continuation value pair if there is a prior history at which the payoff pair is (w, v). Let $b(\nu)$ be the set of all interim value pairs at interim histories arising from σ and α where the agent's interim belief is ν. Let $c(\mu)$ be the set of all continuation value pairs at prior histories arising from σ and α where the agent's belief is μ.

For any σ and α the associated correspondence c is self-generating. In particular

1. $b(\nu) \subset B(c)(\nu)$ for all ν.

2. $c(\mu) \subset C(b)(\mu)$ for all μ.

from which it follows that $c \subset C(B(c))$.

To prove the first claim above, an interim value pair (w, v) belongs to $b(\nu)$ if and only if there is an interim history at which the agent has belief ν and takes an action a such that $(w, v) = (1-\delta)(x, u) + \delta(w', v')$ where $(x, u) = (x(a, \nu), u(a, \nu))$ and (w', v') is the subsequent pair of continuation values. In particular $(x, u) \in \mathcal{U}(\nu)$ and $(w', v') \in c(f(\nu))$. Moreover since α is an optimal strategy for the agent it yields an interim value higher than his minmax value, i.e. $w \geq m(\nu)$. All of the above is equivalent to $(w, v) \in B(c)(\nu)$.

To prove the second claim, a continuation value pair (w, v) belongs to $c(\mu)$ if and only if there is a prior history at which the agent's belief is μ and the principal chooses a lottery over interim beliefs $\tilde{\nu}$ such that $(\mu, w, v) = \mathbf{E}_{\tilde{\nu}}(\nu, \tilde{w}(\nu), \tilde{v}(\nu))$ where for each ν, the interim value pair $(\tilde{w}(\nu), \tilde{v}(\nu))$ belongs to $b(\nu)$. This is equivalent to $(w, v) \in C(b)(\mu)$.

Next we have the following facts

1. The union of self-generating correspondences is self-generating.

2. If W is self-generating then for every prior μ and for every $(w, v) \in C(B(W))(\mu)$ there is a mechanism and an optimal strategy such that (w, v) is the payoff profile.

3. If W is self-generating then so is $C(B(W))$.

The first and third claims are standard and we omit the proofs. For the second claim, let W be a self-generating correspondence. For any $(\mu, w, v) \in \operatorname{graph} C(B(W))$ there is a lottery $\tilde{\nu}$ over interim beliefs such that for each realization ν, there is $(x, u) \in \mathcal{U}(\nu)$ such that

$$(\mu, w, v) = \mathbf{E}_{\tilde{\nu}}(\nu, \tilde{w}(\nu), \tilde{v}(\nu)) \tag{12}$$

where for some $(w', v') \in W(f(\nu))$,

$$(\tilde{w}(\nu), \tilde{v}(\nu)) = (1-\delta)(x, u) + \delta(w', v') \tag{13}$$

and moreover

$$\tilde{w}(\nu) \geq m(\nu). \tag{14}$$

Associated with (x, u) and ν above is an action $a(\nu)$ such that

$$(x, u) = \mathbf{E}_\nu(x(a(\mu), s), u(a(\nu), s)) \tag{15}$$

The mechanism and strategy are defined as follows. The principal uses a direct obfuscation mechanism in which when the agent's belief is μ_t, the principal sends messages according to the associated lottery $\tilde{\nu}$. Then when the agent receives message ν, his strategy specifies that he chooses action $a(\nu)$. The principal and agent behave as above so long as the agent has played the specified action at every history in the past. Otherwise the principal sends an uninformative message (i.e. $\nu = \mu$ with probability 1) and the agent plays his optimal continuation strategy in the no-information mechanism.

First, to show that this strategy is optimal for the agent, observe that any deviation gives the agent his minmax continuation value $m(\nu)$ and by Equation 14 the prescribed strategy gives at least that value. It remains only to show that the mechanism/strategy pair yields the promised values. Fix $(\mu, w, v) \in \text{graph}\, C(B(W))$. The mechanism generates a stochastic process for (x_t, u_t), the flow utilities, and (w_t, v_t) the promised continuation values. By Equation 12 and Equation 13,

$$(w, v) = \mathbf{E}\left[(1-\delta)(x_0, u_0) + \delta(w_1, v_1)\right]$$

and again by Equation 12 and Equation 13,

$$(w, v) = \mathbf{E}\left[(1-\delta)(x_0, u_0) + \delta\left[(1-\delta)(x_1, u_1) + \delta(w_2, v_2)\right]\right]$$

and by induction

$$(w, v) = (1-\delta)\mathbf{E}\left[\sum_{s=0}^{\infty} \delta^s (x_s, u_s)\right]$$

which is the expected discounted payoff from the mechanism/strategy profile beginning with prior belief μ.

Now let W^* be the union of all self-generating correspondences. It follows from the facts above that $W^* = C(B(W^*))$. In particular $W^* \subset C(B(W^*))$ by the first point, and if $W^* \neq C(B(W^*))$, then by the third point $C(B(W^*))$ is a self-generating correspondence that is larger than the union of all self-generating correspondences which is impossible.

Furthermore, for every μ, $W^*(\mu)$ is the set of all payoff pairs at prior belief μ associated with all possible mechanisms and optimal strategies. Thus the function

$$V(\mu, w) := \max\{v : (w, v) \in W^*(\mu)\}$$

gives the maximum payoff to the principal among all mechanisms when the prior belief is μ. We will prove item 3 by showing that V is a fixed point of the functional operator T, in other words that $V = V^*$. First, to establish that V is well-defined, we now show that W^* has compact graph so that the maximum exists.

Consider the following sequence of correspondences.

$$\begin{aligned} W^0(\mu) &= [m(\mu), M(\mu)] \times \left[\min_{a,s} u(a,s), \max_{a,s} u(a,s)\right] \\ W^k &= C(B(W^{k_1})) \end{aligned}$$

The correspondence W^0 has compact graph. Now if a correspondence g has compact graph then so do both $B(g)$ and $C(g)$. In the former case this is because the convex combination of compact-graph correspondences has compact graph. In the latter case this is because by definition graph $C(g)$ is just the convex hull of graph g. Thus, by induction, W^k has compact graph for all k. Moreover, W^0 includes all possible feasible payoffs for principal and agent and therefore $B(W^0) \subset W^0$ and in particular graph $B(W^0) \subset$ graph W^0. It follows that

$$\operatorname{graph} W^1 = \operatorname{graph} C(B(W^0)) \subset \operatorname{graph} C(W^0) = \operatorname{graph} W^0$$

since W^0 has convex graph.

Now if $g \subset g'$ then both graph $C(g) \subset$ graph $C(g')$ and graph $B(g) \subset$ graph $B(g')$ and thus by induction,

$$\operatorname{graph} W^k \subset \operatorname{graph} W^{k-1}$$

for all k. We have thus shown that graph W^k is a nested sequence of compact sets. The infinite intersection $\tilde{W} = \cap_k W^k$ thus has a compact graph which includes the graph of every self-generating correspondence, in particular the graph of W^*.

Indeed we next show that $\tilde{W}$ is itself a self-generating correspondence. Since W^* is the union of all self-generating correspondences, it will then follow that $\tilde{W}$ in fact equals W^* and thus that W^* has compact graph.

Let $(w,v) \in \tilde{W}(\mu)$. Then $(w,v) \in W^k$ for all k. Therefore, for every k there exists four points[31] (not necessarily distinct)

$$(\nu_j^k, w_j^k, v_j^k) \in \operatorname{graph} B(W^{k-1})$$

where $j = 1, \dots, 4$, such that for weights $\alpha^k, j \in [0,1]$ with $\sum_j \alpha_j^k = 1$,

$$(\mu, w, v) = \sum_j \alpha_j^k (\nu_j^k, w_j^k, v_j^k).$$

In addition for all $j = 1, \dots, 4$, there exist

$$\begin{aligned}(x_j^k, u_j^k) &\in \mathcal{U}(\nu_j^k) \\ (\tilde{w}_j^k, \tilde{v}_j^k) &\in W^{k-1}(f(\nu_j^k))\end{aligned}$$

such that

$$(w_j^k, v_j^k) = (1-\delta)(x_j^k, u_j^k) + \delta(\tilde{w}_j^k, \tilde{v}_j^k)$$

and $w_j^k \geq m(\nu_j^k)$. The sequence (in k) $\{\alpha_j^k, (\nu_j^k, w_j^k, v_j^k), (u_j^k, u_j^k), (\tilde{w}_j^k, \tilde{v}_j^k)\}$ has a convergent subsequence whose limit $\{\alpha_j, (\nu_j, w_j, v_j), (u_j, u_j), (\tilde{w}_j, \tilde{v}_j)\}$ satisfies

$$\begin{aligned}(\nu_j, w_j, v_j) &\in \operatorname{graph} \tilde{W} \\ (\mu, w, v) &= \sum_j \alpha_j (\nu_j, w_j, v_j) \\ (x_j, u_j) &\in \mathcal{U}(\nu_j) \\ (\tilde{w}_j, \tilde{v}_j) &\in \tilde{W}(f(\nu_j)) \\ (w_j, v_j) &= (1-\delta)(x_j, u_j) + \delta(\tilde{w}_j, \tilde{v}_j) \\ w_j &\geq m(\nu_j)\end{aligned}$$

[31] By Caratheodory's theorem. Since $W^k = C(B(W^{k-1}))$, the graph of W^k is equal to the convex hull of the graph of $B(W^{k-1})$. The latter is a subset of $\mathbf{R}^3$ and hence any element of the graph of W^k can be written as a convex combination of four points in graph$(B(W^{k-1}))$.

Thus $(\mu, w, v) \in \operatorname{graph} C(B(\tilde{W}))$, i.e. $(w, v) \in C(B(\tilde{W}))(\mu)$ and we have shown that

$$\tilde{W} \subset C(B(\tilde{W})).$$

The proof of item 3 is now concluded by showing that V is a fixed point of T. It will follow from the next two claims.

1. If $v = TV(\mu, w)$ then $(w, v) \in W^*(\mu)$.
2. If $(w, v') \in W^*(\mu)$ then $v' \leq TV(\mu, w)$.

Define the correspondence

$$D = (1 - \delta)\mathcal{U} + \delta(V^\mu \circ f).$$

Since $V^\mu \subset W^*$, we have $D \subset (1 - \delta)\mathcal{U} + \delta(W^* \circ f)$. Note that $(\nu, w, v) \in \operatorname{graph} D^{\mathcal{W}}$ implies both $(\nu, w, v) \in \operatorname{graph} D$ and $w \geq m(\nu)$. These facts in turn imply that $(\nu, w, v) \in \operatorname{graph} B(W^*)$. Thus

$$\operatorname{graph} D^{\mathcal{W}} \subset \operatorname{graph} B(W^*).$$

Now let $v = TV(\mu, w)$. Since $TV = \operatorname{cav} D^{\mathcal{W}}$, there exists a lottery $(\tilde{\nu}, \tilde{w})$ in $\mathcal{W}$ and a selection $\bar{v}(\tilde{\nu}, \tilde{w}) \in D^{\mathcal{W}}(\tilde{\nu}, \tilde{w})$ such that $\mathbf{E}(\tilde{\nu}, \tilde{w}, \bar{v}) = (\nu, w, v)$. The lottery $(\tilde{\nu}, \tilde{w}, \bar{v}(\tilde{\nu}, \tilde{w}))$ is over the graph of $D^{\mathcal{W}}$ which is a subset of the graph of $B(W^*)$. Thus by the definition of C, we have shown $(w, v) \in C(B(W^*))(\mu)$. Since $W^* = C(B(W^*))$, we have established the first claim above.

Since $V(\nu, w) = \max\{v : (w, v) \in W^*(\nu)\}$, for every $(\nu, w, v) \in \operatorname{graph} B(W^*)$ there is a $v' \geq v$ such that $v' \in D^{\mathcal{W}}(\nu, w)$. If $(w, v) \in W^*(\mu) = C(B(W^*))(\mu)$ then by the definition of C, there exists a lottery $(\tilde{\nu}, \tilde{w}, \tilde{v})$ on graph $B(W^*)$ such that $\mathbf{E}(\tilde{\nu}, \tilde{w}, \tilde{v}) = (\mu, w, v)$. The lottery $(\tilde{\nu}, \tilde{w})$ is a lottery on $\mathcal{W}$, and for every realization (ν, w, v) we can replace v with an element $v'(\nu, w) \in D^{\mathcal{W}}(\nu, w)$ such that $v'(\nu, w) \geq v$. We then obtain a lottery $(\tilde{\nu}, \tilde{w}, v'(\tilde{\nu}, \tilde{w}))$ on the graph of $D^{\mathcal{W}}$ such that the expected value of $(\tilde{\nu}, \tilde{w})$ is (μ, w) and

$$\mathbf{E}v'(\tilde{\nu}, \tilde{w}) \geq \mathbf{E}\tilde{v} = v.$$

Since $TV = \operatorname{cav} D^{\mathcal{W}}$, and thus $TV(\mu, w)$ is the maximum expectation over all such lotteries, it is no smaller than $\mathbf{E}v'(\tilde{\nu}, \tilde{w})$ and we have established the second claim.

□

D.1 The Concavification of a Correspondence

Let $g : \mathcal{W} \rightrightarrows \mathbf{R}$ be a correspondence whose domain $\mathcal{W}$ is a subset of $\mathbf{R}^n$. We defined the concavification $\operatorname{cav} g$ to be the smallest concave function that is pointwise no-smaller than any element of the image of g. Let $\operatorname{hypo} g$ be the hypograph of g:

$$\operatorname{hypo} g = \left\{ (x, y) \in \mathbf{R}^{n+1} : \exists y' \in g(x) \text{ such that } y \leq y' \right\}$$

Let F be the convex hull of $\operatorname{hypo} g$ and consider the function h defined by

$$h(x) = \sup \{ y : (x, y) \in F \} .$$

The function h is concave (see Rockafellar (1997)Theorem 5.3). Moreover it is pointwise no smaller than any element in the image of g. Indeed it is the pointwise smallest such concave function and therefore it equals $\operatorname{cav} g$.

Now $(x, y) \in F$ if and only if there exists a lottery $(\tilde{x}, \tilde{y})$ over $\operatorname{hypo} g$ such that $\mathbf{E}(\tilde{x}, \tilde{y}) = (x, y)$. Thus

$$\begin{aligned} h(x) &= \sup_{\{(\tilde{x},\tilde{y}) : \mathbf{E}\tilde{x} = x, \exists y \in g(\tilde{x}) \text{ s.t. } y \geq \tilde{y}\}} \mathbf{E}\tilde{y} \\ &= \sup_{\{(\tilde{x},\tilde{y}) : \mathbf{E}\tilde{x} = x, \tilde{y} \in g(\tilde{x})\}} \mathbf{E}\tilde{y} \\ &= \sup_{\{(\tilde{x},\tilde{y}) \in \operatorname{graph} g : \mathbf{E}\tilde{x} = x\}} \mathbf{E}\tilde{y} \end{aligned}$$

Finally, note that by the Theorem of the Maximum, if g has compact graph then so does $\operatorname{cav} g$, and in particular we can replace the sup with max above.

D.2 Deriving The Value Function For The Work-Shirk Example

We begin with a proof of Lemma 4.

Proof of Lemma 4. The following expression gives the agent's minmax value at any $\mu \geq 1/2$. With no information about the state and a belief $\mu \geq 1/2$ the agent will shirk for the remainder of the process. With probability μ the state is 1 and will remain so and he will earn 1 forever. With the remaining probability the state is 0 and he will earn -1 until the state changes after which he will earn 1.

$$\begin{aligned} m(\mu) &= 1 - 2(1-\mu)\sum_{t=0}^{\infty}\delta^t(1-\Lambda)^t \\ &= 1 - 2(1-\mu)\frac{1-\delta}{1-\delta(1-\Lambda)} \end{aligned}$$

Turning to the static optimal mechanism, the agent's value at $\mu = 1/2$ can be expressed recursively as follows

$$W(1/2) = \delta\left[qW(1/2) + (1-q)\right]$$

because the expected flow payoff from working at belief 1/2 is zero and in the following period the agent will either move back to $\mu = 1/2$ or to $\mu = 1$ and earn 1 forever. The probability of returning to $\mu = 1/2$ is

$$q = \frac{1 - f(1/2)}{1 - 1/2} = 2(1 - f(1/2))$$

and $f(1/2) = 1/2 + \Lambda/2$, therefore $q = 1 - \Lambda$ and

$$\begin{aligned} W(1/2) &= \frac{\delta\Lambda}{1-\delta(1-\Lambda)} \\ &= \frac{(1-\delta) + \delta\Lambda - (1-\delta)}{1-\delta(1-\Lambda)} \\ &= 1 - \frac{1-\delta}{1-\delta(1-\Lambda)} \end{aligned}$$

which equals $m(1/2)$. Furthermore, for any $\mu \in [1/2, 1]$,

$$W(\mu) = 2(1-\mu)W(1/2) + (1 - 2(1-\mu))$$

which, like $m(\mu)$ is linear in μ and equals 1 at $\mu = 1$, therefore we have shown that the minmax values and static optimum values for the agent coincide for $\mu \geq 1/2$. That they also coincide for $\mu \leq 1/2$ follows immediately because in both mechanisms the agent is working and obtaining no information until his belief crosses $\mu = 1/2$ whereupon he earns the corresponding continuation value. □

Proof of Proposition 2. The analysis will be based on the graph of the value function given in the statement of the proposition.

$$V(\mu, w) = \alpha \bar{v}(\mu) + (1-\alpha)\underline{v}(\mu) \tag{16}$$

To begin with, we record some important facts about the static-optimum and full-information values, i.e. the points $(\mu, m(\mu), \bar{v}(\mu))$ (applying here Lemma 4), and $(\mu, M(\mu), \underline{v}(\mu))$ in the graph of V. Recall that in the static-optimal mechanism, the agent shirks if and only if he attaches probability 1 to state 1. Thus the agent obtains flow payoff of 1 in any period he shirks, and moreover he obtains a flow payoff of 0 whenever he works. On the other hand, the principal obtains a flow payoff of 1 whenever the agent works and 0 whenever the agent shirks. Thus, in the static-optimal mechanism, the two parties' values sum to 1, i.e. $m(\mu) + \bar{v}(\mu) = 1$ for all μ.

Indeed the same is true for the full-information mechanism since again the agent works if and only if he attaches probability 1 to state 1. Thus $M(\mu) + \underline{v}(\mu) = 1$ for all μ. It then follows from the linearity in Equation 16 that the graph of V is a subset of the plane $\mathcal{L} \subset [0,1] \times \mathbf{R} \times \mathbf{R}$ defined by $v + w = 1$. Therefore, the graph of $V^\mu \circ f$ is also a subset of $\mathcal{L}$. Now consider the graph of

$$(1-\delta)\mathcal{U} + \delta(V^\mu \circ f).$$

Since the agent has two actions, $\mathcal{U}$ always has two points in its image and therefore the above is the union of two sets:

$$\text{graph}\,[(1-\delta)\mathcal{U}_{\texttt{work}} + \delta(V^\mu \circ f)] \cup \text{graph}\,[(1-\delta)\mathcal{U}_{\texttt{shirk}} + \delta(V^\mu \circ f)]$$

where $\mathcal{U}_{\texttt{work}}$ and $\mathcal{U}_{\texttt{shirk}}$ are the functions giving the flow utility vectors associated with shirk and work, namely

$$\mathcal{U}_{\texttt{work}}(\mu) \equiv (0,1)$$

and

$$\mathcal{U}_{\texttt{shirk}}(\mu) = (2\mu - 1, 0).$$

Since the graph of $\mathcal{U}_{\texttt{work}}$ also belongs to $\mathcal{L}$, we know that

$$\text{graph}\,[(1-\delta)\mathcal{U}_{\texttt{work}} + \delta(V^\mu \circ f)] \subset \mathcal{L}.$$

Moreover the graph of $\mathcal{U}_{\texttt{shirk}}$ is always weakly below (in terms of the principal's payoff) $\mathcal{L}$, and intersects that plane in exactly one point, $(\mu, w, v) = (1, 1, 0)$. Therefore, when we consider the concavification

$$TV = \text{cav}\left[(1-\delta)\mathcal{U} + \delta(V^{\mu} \circ f)\right]^{\mathcal{W}},$$

we know that its graph must be weakly below $\mathcal{L}$. (Because the linear function on $\mathcal{W}$ whose graph is contained in $\mathcal{L}$ (namely V) is a concave function no smaller than the mapping being concavified and TV is defined as the pointwise smallest such function.) I will now show that its graph in fact coincides with $\mathcal{L}$ over the domain $\mathcal{W}$, and therefore that the graph of TV coincides with the graph of V which will prove that V is the fixed point of T.

To that end, we take account of some subsets of

$$\mathcal{K} = \mathcal{L} \cap \text{graph}\left[(1-\delta)\mathcal{U} + \delta(V^{\mu} \circ f)\right].$$

First, we have already observed that $(1, 1, 0) \in \mathcal{K}$. Next consider the point $(0, M(0), \underline{v}(0))$. This point corresponds to the full-information mechanism when the agent begins with degenerate belief $\mu = 0$. With a degenerate belief there is no information to give so these values are generated by simply having the agent work and then giving both parties their continuation values at $f(0)$, i.e.

$$(M(0), \underline{v}(0)) = (1-\delta)(0, 1) + \delta(M(f(0)), \underline{v}(f(0)))$$

hence $(0, M(0), \underline{v}(0)) \in \mathcal{K}$.

Finally, consider points of the form $(\mu, m(\mu), \bar{v}(\mu))$ for $\mu \leq 1/2$. These correspond to the no-information mechanism. In the no-information mechanism when $\mu \leq 1/2$, the agent works and his beliefs are updated to $f(\mu)$. Hence

$$(m(\mu), \bar{v}(\mu)) = (1-\delta)(0, 1) + \delta(m(f(\mu)), \underline{v}(f(\mu)))$$

so that these points all belong to $\mathcal{K}$.

The convex hull of the subset we have identified is a subset of $\mathcal{L}$ and indeed equals the graph of V.Therefore the concavification TV must be weakly above $\mathcal{L}$. Since we have already shown it is weakly below $\mathcal{L}$, the proof is complete. □