Beeps*

Jeffrey C. Ely †

December 15, 2013

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. *Keywords: beeps, obfuscation principle* .

*Thanks especially to Toomas Hinnosaar and Kevin Bryan for early discussions on this topic. I received excellent questions, comments, and suggestions from Johannes Horner, Moritz Meyer-ter-Vehn, Simon Board, Sergiu Hart and Debraj Ray. 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 an 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 experimentation.

[1] Cite Orlov.

[2] Cite Che-Horner, Gershkov-Kremer-Perry

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 **?**. 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 **?**, 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

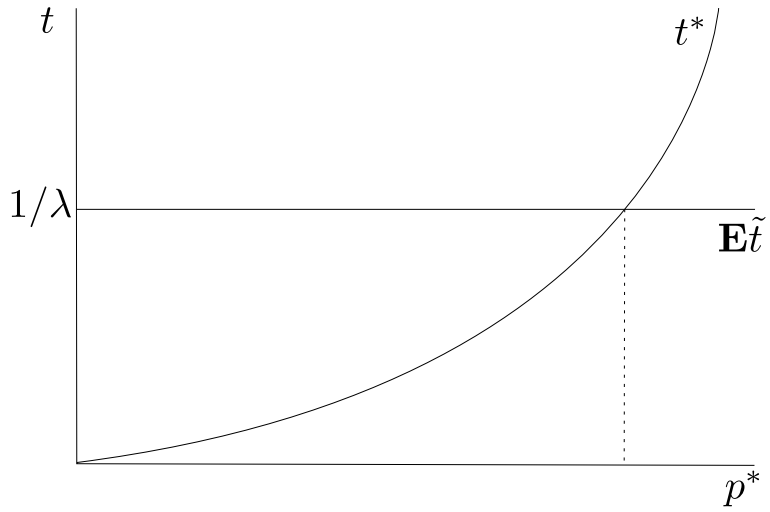
$$t^* = -\frac{log(1-p^*)}{\lambda}$$

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

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

[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.

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 before she can't resist.[6]

[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.

[6]The precise turning point is the threshold that satisfies $1 = log(1 - p^*)$ which is $1 - 1/e$, roughly .63.

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

[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.

To characterize the optimal mechanism it would be intractable to optimize over the enourmous set of feasible information policies. Fortunately, building on ideas from **?**, **?**, **?**, we can appeal to the *obfuscation principle* and capture the full set of feasible policies by a tractable family of simple, *direct obfuscation mechanisms*. 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.

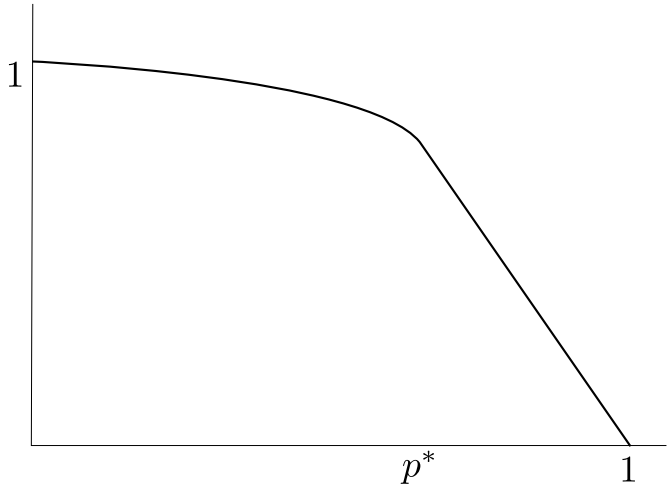
1. $\mathbf{E}(\nu_t \mid \mu_t) = \mu_t$,
2. $\dot{\nu}_t = \lambda\,(1 - \nu_t)$.

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.[8]

This enables us to reformulate the problem and solve it using dynamic programming, using μ_t as a state variable and explicitly incorporating the

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

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 **?** and **?**. 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 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 **?**. 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.[9] 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.[10] 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 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

[9] 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.

[10] Toomas Hinnosaar first suggested this policy.

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.

Following a literature review below, in the remainder of this paper I outline the general principal-agent model with arbitrary Markov process, action space, and payoffs of which email beeps is a special case. I show a series of geometric operations that derive the optimal value function for the principal and I show how to infer from it the optimal policy. Each of the problems I discuss have an analogous static model and I discuss the connection between the dynamic optimum and the static optimum characterized by **?**. I then consider various extensions and discuss ongoing work.

1.1 Related Literature

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

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

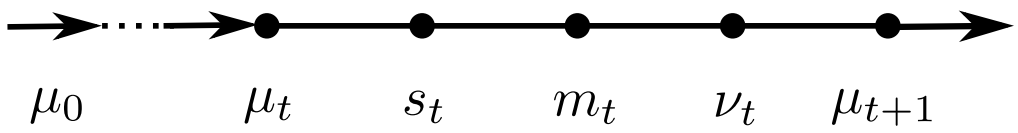
where $M_\mu = \sum_s \mu(s) M_s$.

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 v:

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

The principal's payoff $u(a)$ also depends on the agent's action and therefore indirectly depends on the agent's interim belief $u(\nu) = u(a(\nu))$. Indeed we can take the 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

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

is a bounded upper semi-continuous payoff function for the principal[11] and that the principal is maximizing the expected value of his long-run average payoff

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

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 one in which instead of choosing a policy, the principal is directly specifying a stochastic process for the agent's beliefs.[12]

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

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

[11] When the agent is maximizing the expected value of v, 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.

[12] 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.

Conversely any stochastic process with initial belief μ_0 satisfying these 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 **??**. 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 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 2 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 **?** and **?**, 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 larger than the function being concavified.[13] We obtain the following functional equation.

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

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 value function. By standard arguments this operator is a contraction and therefore has a unique fixed point which can be found by iteration. 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:

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

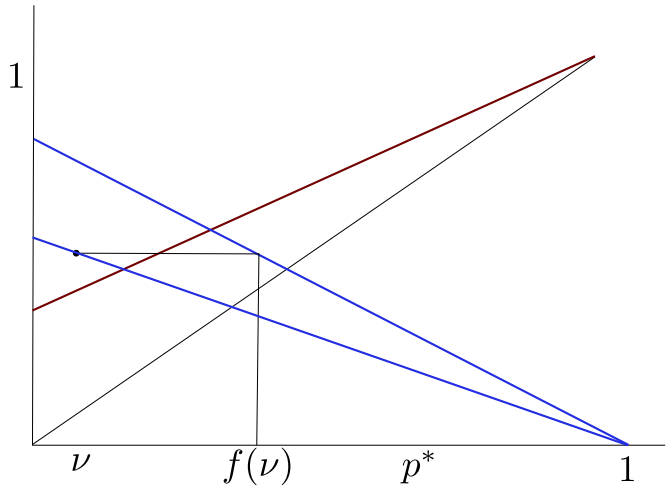
The principal's indirect utility function is

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

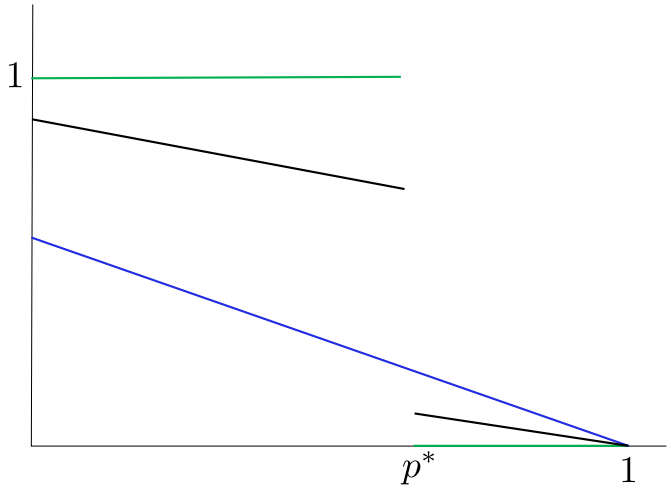
[13] Equivalently it is the pointwise infimum of all pointwise larger concave functions. Geometrically, one can identify a function with its epigraph. Then the concavification is the epigraph that is obtained by the convex hull of the original epigraph. Rockafeller calls it... FXIME

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

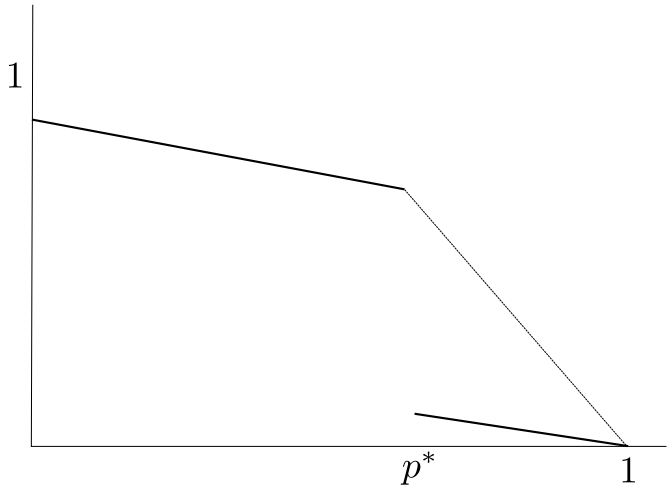
In the appendix I derive the value function and optimal policy analytically. Here we will follow a sequence of diagrams to visualize the derivation and gain intuition. Refer to the Bellman equation in Equation 2. Consider as an initial guess, a linear V. We can trace through the operations on the right-hand 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.[14]

[14]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

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 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.[15] The optimal value function is represented below.

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.

---

does not, then his belief jumps to 0 and beep-on continues from there. Thus, the value at μ_0 is just $(1 - \mu_0)V(0)$.

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

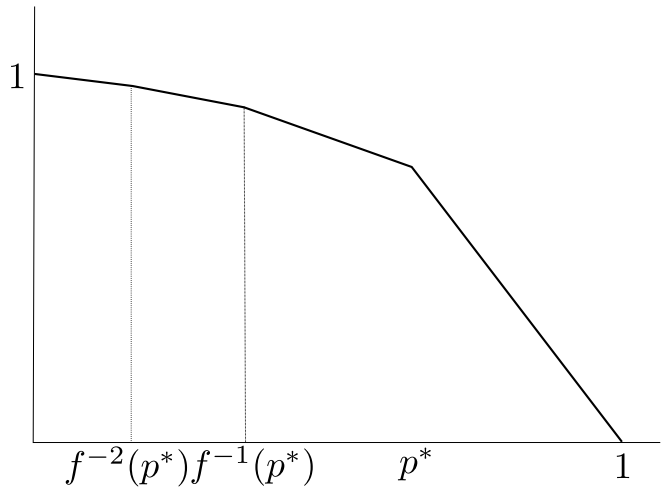
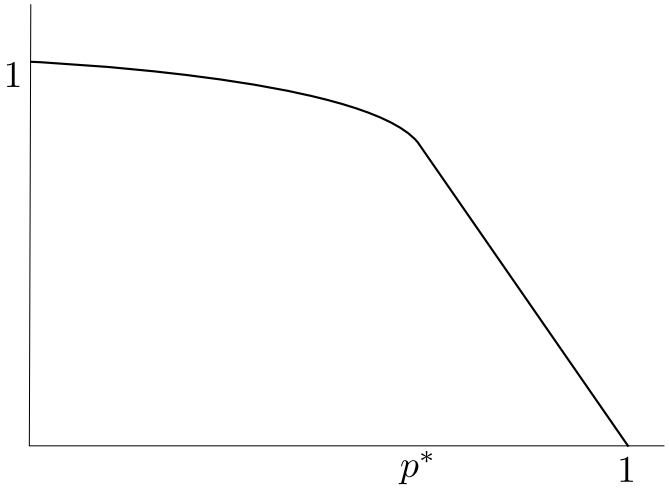


Figure 1: The optimal value function for the Beeps problem.



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 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

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

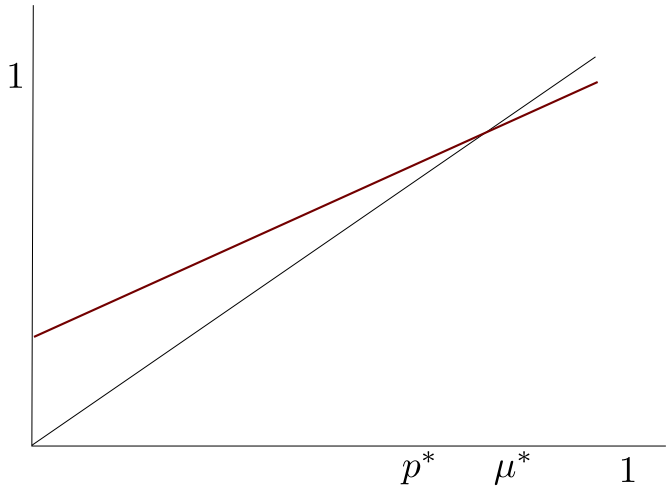
Notice that $r/(r+\lambda)$ is the discounted average value of receiving a flow payoff of 1 until a termination date which arrives at Poisson rate λ. Indeed that is exactly the *initial* (i.e. starting at $\mu = 0$, not at $\mu = p^*$) discounted value from the beep-on policy. In the optimal mechanism 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.

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 $\mu^* > p^*$.[16]

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

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

[16] 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.

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 $\mathcal{I}$.)

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.

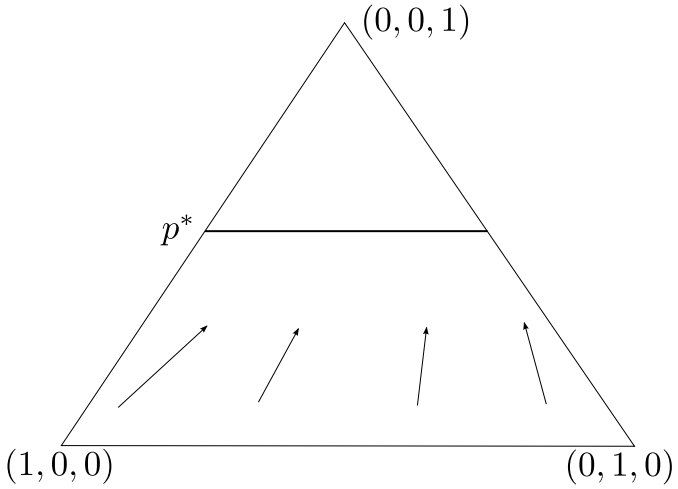
In particular, the value $V(1)$ is positive now because the agent will periodically switch back to working when his beliefs jump down to p^*.

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.[17]

[17]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

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.

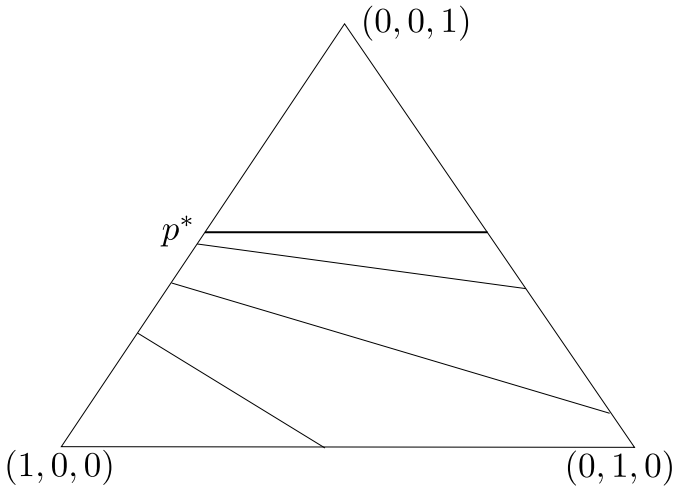


This analysis of this problem follows similar lines as the two-state email beep problem. In the diagram below the lines represent points which lead by iterations of f to the threshold line. If we begin with a candidate value function which is linear and equal to zero at the $s = 2$ vertex, iterations lead to a piecewise linear value function with kinks along these segments.[18] The continuous time limit value function will therefore be linear along these line segments but strictly concave along rays toward the

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. Jerome Renault, Eilon Solan, Nicolas Vielle, in concurrent work are analyzing the three state problem in more detail (private communication.)

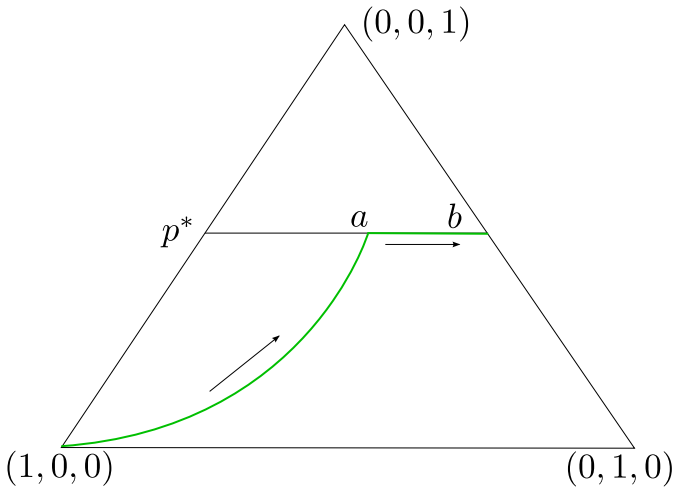
[18] These are not the level sets. As we will show below, the value declines as we move to the right along the p^* line. This pattern will thus be preserved at all of its inverse images under f.

$s = 2$ 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.



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 we will show in the next section, the dynamic optimum 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 intitial 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.

Consider the following indirect utility function with 3 steps:

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

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: 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 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:

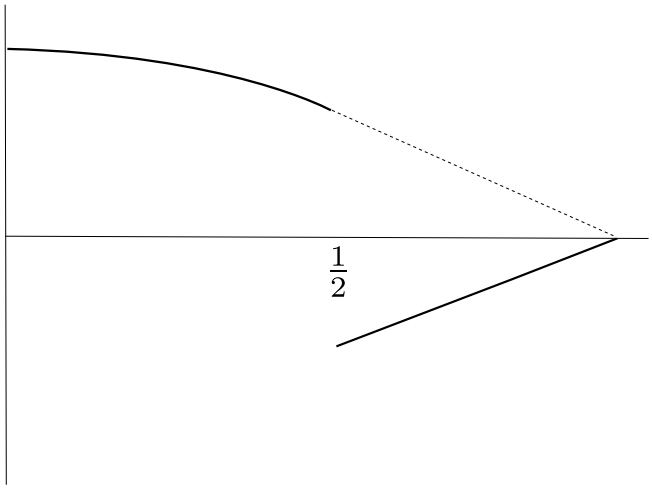
$$\frac{dW}{dt}(t) = e^{-(t^*-t)}\left[W(t^*) - 1\right]$$

(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

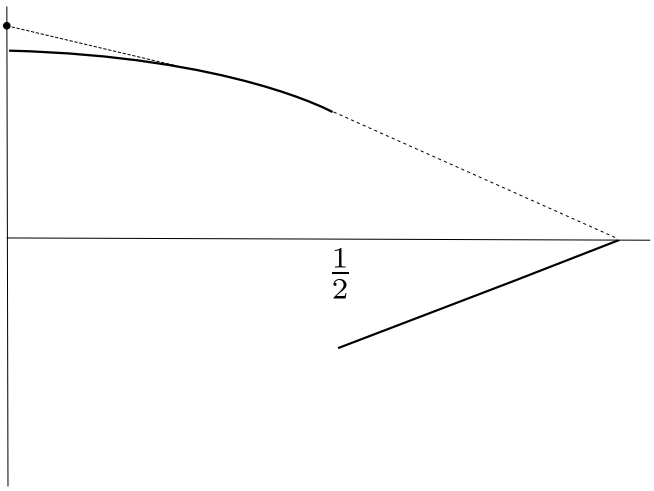
$$(1 - e^{-t(\mu)}) + e^{-t(\mu)}V(p^*)$$

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 was 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.

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

4 Connection With Static Problems

5 Extensions

5.1 Long-Run/Strategic Agent

5.2 Principal's payoff depends on the state

When the principal's payoff depends on the state s as well as the agent's action a,

$$u(a, s)$$

the nature of the incentive problem changes but it can be analyzed without a change to the basic structure. Indeed we can express the indirect utility function as follows

$$u(\nu) = \mathbf{E}_\nu u(a(\nu), s) \tag{5}$$

since the agent's preferred action is still a function of his belief $a(\nu)$ and conditional on the agent's belief being ν it follows that the probability distribution over states s is exactly ν. This latter point deserves emphasis because of course the principal knows the true state and only the agent has beliefs ν. Indeed the principal's preferences over the agent's beliefs are s-dependent since his preferences over a are s-dependent. However, an immediate extension of the ideas[19] behind the obfuscation principle justifies the reduced-form above.

Lemma 3. *All policies that generate a given stochastic process (μ_t, ν_t) generate the same stochastic process over the larger space (μ_t, ν_t, s_t). Indeed, in any t for any state s, and for any measurable set of interim beliefs $B \in \Delta(S)$,*

$$\text{Prob}(B \times \{s\} \mid \mu_t) = \text{Prob}(B \mid \mu_t) \int_{\nu \in B} \nu(s) d\sigma_{\mu_t}(B)$$

where σ_{μ_t} is the total probability distribution over ν_t induced by the policy at a time t history in which the agent's belief is μ_t.

This lemma is immediate because $\nu(s)$ is the conditional probability of s given a history which leads the agent to interim belief ν. It implies that the principal chooses directly a stochastic process for (μ_t, ν_t, s_t) satisfying

[19] See the discussion following Lemma 1.

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

and since the third condition identifies exactly one extended stochastic process associated with each process on the smaller domain (μ_t, ν_t), the feasible set for the principal is unchanged: the set of processes satisfying the first two conditions.

We can thus express the principal's optimization problem exactly as before, noting only that the indirect utility function u is now derived as in Equation 5.

5.3 Actions affect states

5.4 Actions generate information

5.5 Principal is learning over time

A The Obfuscation Principle

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

$$\sigma_s(B) = \int_{\nu \in M} \frac{\nu(s)}{\mu_t(s)} dZ(\nu). \tag{6}$$

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 6 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. FIXME CITE BILLINGSLEY Thus, $y(\nu) = \nu(s)$ is the Bayesian posterior probability of state s upon receiving the message ν as desired.

B Optimal Mechanism For The Beeps Example

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 $\alpha \in (0,1)$,

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

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 $p^{**} = f(p^*)$,

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

implying that

$$W(p^{**}) = \frac{(1-\delta)(1-p^{**})}{1-p^*-\delta(1-p^{**})} \tag{7}$$

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

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

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

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 7 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.[20] To see this, note that the latter value is equal to the expected discounted waiting

[20]Multiplying by r in the formulas below normalizes payoffs in the same way that multiplying by $(1-\delta)$ does in discrete time.

time before the first email arrival, i.e.

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

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 is 1 until time t^*, after which he obtains continuation value $V(p^*)$. This yields discounted expected value

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

And since

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

we have

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

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

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