

Dynamic Amnesty Programs

Sam Kapon*

[\[Click here for the latest version\]](#)

16th February 2021

Abstract

A regulator faces a stream of agents each engaged in crime with stochastic returns. The regulator designs an amnesty program, committing to a time path of penalty reductions for criminals who self-report before they are detected. In an optimal time path, the intertemporal variation in the returns from crime can generate intertemporal variation in the generosity of amnesty. I construct an optimal time path and show that it exhibits amnesty cycles. Amnesty becomes increasingly generous over time until it hits a bound, at which point the cycle resets. Agents engaged in high return crime self-report at the end of each cycle, while agents engaged in low return crime self-report always. An extension to multi-agent organizations, like price-fixing cartels, is examined, where the role for time variation in amnesties is magnified by a preemption effect.

Keywords— Dynamic Mechanism Design, Self-Reporting, Amnesty, Cartels, Crime

1. Introduction

To stop ongoing crime, a regulator can offer preferable treatment to criminals who self-report. These *amnesty*, or *self-reporting programs*, appear in such diverse contexts as illegal gun

*New York University, sdk301@nyu.edu. I am indebted to Sylvain Chassang for his guidance and patience. I am grateful to Guillaume Fréchette and Dilip Abreu for their support. This paper has benefited from discussions with and comments from David Pearce, Erik Madsen, Debraj Ray, Paula Onuchic, Dmitry Sorokin, Mauricio Ribiero, Alessandro Lizzeri, Nikhil Vellodi, Andrew Schotter, Stefan Bucher, Lester Chan, Ariel Rubinstein, Maher Said, Dhruva Bhaskar and Alistair Wright. I thank participants at the 2020 Young Economist Symposium and various NYU seminars for their comments.

ownership, collusion, desertion in war, tax evasion, espionage¹, civil conflict, and corruption. For instance, the U.S. Department of Justice operates a leniency program for self-reporters, which has become its “most important investigative tool for detecting cartel activity.”² The Red Army’s amnesty for military desertion in June 1919 induced the return of over 100,000 deserters.³ Australia’s gun buy-back of 1997 collected more than 650,000 weapons.⁴ The Chieu Hoi program offered amnesty to defectors during the Vietnam war, enticing over 100,000.⁵

An extensive theoretical literature has investigated the use of self-reporting programs in one-shot regulation.⁶ Less attention has been paid to the inter-temporal properties of these programs, which are often offered on a repeated, time-limited, basis. The Red Army’s Central Anti-Desertion Commission operated repeated amnesty periods, interspersed with periods of harsh enforcement (Wright, 2012) and a similar program was applied to desertion in French militaries in the 17th and 18th centuries (Forrest et al., 1989). The Brazilian gun buyback program has been run four times since 2013 (Macinko et al., 2007). The U.S. has operated a number of tax-related self-reporting programs, often on a repeated, time-limited, basis.⁷ Other programs are offered continuously. For instance, the U.S. Department of Justice’s cartel leniency program and the Mexican gun buyback are, and the Chieu Hoi program was, run continuously without explicit adjustment to the terms of self-reporting (Wosepkat, 1971).

In this paper, I ask: how should the terms of self-reporting programs be designed over time? Should programs be permanent, with constant terms – i.e. static – or, should they fluctuate inter-temporally – i.e. dynamic? I study a mechanism design problem in which criminal agents arrive randomly and their returns from crime are private, idiosyncratic and evolve over time. A regulator commits to a time-path of penalty reductions for agents who self-report before they are detected that applies uniformly to all agents. The range of possible punishments is bounded and the agent may be exogenously detected, at which point the regulator applies the maximum punishment possible. The agents’ only decision is when, if ever, to self-report.

The inter-temporal variation in the returns from crime drives the desirability of dynamic programs for the regulator in this environment. If an agent’s returns from crime can transition from a high to low state, static self-reporting programs are subject to exploitation by the agents. The intuition for this is best seen by comparing two extreme policies. The first

¹Such as the amnesties offered to British informants by the Irish Republican Army in the 1980’s.

²<https://www.justice.gov/atr/leniency-program>

³See Figes (1990).

⁴See Leigh and Neill (2010).

⁵See Wosepkat (1971)

⁶See for instance, Kaplow and Shavell (1994), Malik (1993) and Andreoni (1991) for early contributions.

⁷OECD (2015), Luitel and Sobel (2007)

is a *static program*, offering the same terms for self-reporting at all times. This program lets the agent benefit both from crime while his return is high and from self-reporting once his return is low. At the opposite extreme is a *once-and-for-all program*, in which agents only have one chance to self-report for favorable treatment and are otherwise treated harshly, as if detected exogenously. Under the once-and-for-all program, agents with high returns from crime choose to self-report rather than wait for their return to become low, knowing that by then the option to self-report will be gone. The once-and-for-all program is therefore able to generate self-reporting by higher return agents than the static program.

The drawback of the once-and-for-all program is that agents who arrive after the once-and-for-all opportunity never self-report. The regulator must then balance two forces: (i) enticing contemporaneous agents to self-report by offering a future with less opportunity for self-reporting and (ii) enticing future agents to self-report by not completely shutting down these opportunities, as is done in the once-and-for-all program. This trade-off is explored in the remainder of the paper.

The basic features of the model are motivated by the following observations. First, the returns from crime often accrue slowly over time. Second, returns from crime are private, idiosyncratic and *change over time*. Military deserters face uncertain food and shelter availability, and an uncertain risk of being caught (Forrest et al., 1989). Illegal gun owners may leave crime (Willmer, 1971) or find themselves in need of the money from a gun buyback (Dreyfus et al., 2008). Cartels face fluctuating demand conditions, new entrants de-stabilize cooperation, the risk of detection changes over time (Connor (2007), Gärtner (2014)), and these may be difficult to observe until long after the cartel has been detected, or ever. Third, in many settings of interest, crime has long-term, irreversible effects: a deserter cannot stop being a deserter without military permission, a change in tax payment can spark IRS scrutiny, and in general the evidence of crime is persistent. This irreversibility motivates the assumption that the only way to leave crime is to self-report to the regulator. Finally, amnesty typically takes the form of a reduction in penalties for any agent who self-reports at a given time, motivating the regulator's problem as a choice of a time path for amnesty that applies uniformly to all agents.⁸

The main result (Theorem 2) characterizes an optimal amnesty policy and shows that it takes a *cyclical* form; at regularly spaced times, the regulator offers the lower bound on self-reporting penalties, and in between offers a decreasing schedule of penalties which eventually hits the lower bound and resets the cycle. The decreasing schedule of penalties induces agents with a low return from crime to immediately report, and is chosen so that these agents are indifferent between immediately reporting and waiting until the end of the cycle to report. Agents in the high return state report at the end of each cycle, when the

⁸In Section 5, the features of the model are further discussed and interpreted.

self-reporting penalty is at its lower bound. The frequency of cycles increases with the risk of detection, the penalty for detection, and the rate of transition from high to low return crime.

A *backloading* motive on the part of the regulator is a driving force behind the optimal arrangement of reporting by high return agents. Although exogenous time discounting is equal for the agents and the regulator, agents may be detected and forced to stop committing their crime. This acts like additional time discounting, effectively making the regulator more patient than the agents. Intuitively, the relative patience of the regulator leads to the following backloading result: from the perspective of any time at which high return agents report, the regulator would always prefer to increase the *delay* to the next reporting time of high return agents, if it allows her to increase the *generosity* of the amnesty at that time. This delay is increased until the generosity of amnesty cannot be increased further i.e. the self-reporting penalty hits its lower bound. Repetition of this intuition leads to the optimality of regularly spaced reporting by high return agents, with the minimum self-reporting penalty offered at these times.⁹

In Section 4, I discuss a number of settings in which dynamic self-reporting policies play a role. In the case of military desertion, I recount qualitative evidence from a case study of the Red Army's anti-desertion campaign in Karelia detailed in Wright (2012), among other sources, to argue that the dynamic policies observed were a result of deliberate design intended to induce fast self-reporting. I move to illegal gun ownership and consider how the design results of the model may be applied to improve amnesty and buyback programs. Last, I discuss the application of the model to voluntary disclosure and amnesty programs in tax collection.¹⁰

I extend the model and the intuition to the case of multi-agent crime organizations, where some or all members of the organization can self-report.¹¹ The literature studying such settings has shown that self-reporting programs can be especially powerful when they apply different punishments to different members of an organization, either treating most favorably an agent who reports earliest or using some other features of the relationship between group members to adjust punishments when the crime is reported. I show that there is scope for dynamic amnesty programs to improve the regulator's outcomes relative to static programs, through the former channel, under some assumptions about the tools at

⁹In a setting in which a principal repeatedly supplies an agent with an input to production, Krasikov et al. (2019) similarly finds that backloading incentives that meet individual rationality constraints can generate outcomes that cycle (distortions to efficient allocation that cycle).

¹⁰In some settings, in particular tax collection, the perceived fairness of enforcement may lead to a moral obligation to comply that generates higher compliance than would be implied by enforcement strength and financial considerations alone. In such cases, amnesty may backfire and lead to a deterioration of this moral obligation. I discuss this further in Section 5.

¹¹See for instance Motta and Polo (2003) and Spagnolo (2000) for early work in the case of cartels.

the regulator’s disposal which I detail in Section 6.2. I demonstrate one particular cyclical policy that performs well. Under agent-preferred equilibrium selection, this policy is optimal for the regulator.

After reviewing the literature, I introduce the model in Section 2 and analyze it in Section 3. In Section 4, I present applications of the model. The assumptions of the model and alternative modeling choices are discussed in Section 5. In Section 6, I extend the model to the case of multi-agent organizations. I conclude in Section 7.

Literature. This paper is related to the theoretical literature on self-reporting programs in single and multi-agent settings and the dynamic mechanism design literature, in particular intertemporal price discrimination in the economics and operations research literatures.

The early work of Kaplow and Shavell (1994), Malik (1993), and Andreoni (1991), studied law enforcement and self-reporting behavior in one-shot settings. Much of the subsequent literature is concerned with one-shot self-reporting settings in which the optimal inter-temporal use of amnesties cannot be studied. Nevertheless, the dynamic properties of self-reporting programs have received some attention in the theoretical and empirical literature, although no theory has been developed that accounts for time-variation in the returns from crime. For instance, Marchese and Cassone (2000) rationalizes repeated tax amnesties as a method of discriminating between tax payers who are ex-ante different. Wang et al. (2016) studies how a regulator should design remediation and inspection policies for environmental hazards that arrive randomly over time e.g. leaks. A firm has an option to delay repair of its environmental hazard and the paper focuses on the interaction between the inspection policy and penalties. As the paper shows, when the rate of inspection (which is like the rate of detection in this paper) cannot be chosen but is instead Poisson, optimal self-reporting programs are always static. I focus on the role that dynamic self-reporting programs play absent any control of inspection policies but in the presence of dynamic returns from crime. In this sense, the papers are complementary.¹²

This paper is related to work in dynamic mechanism design such as Battaglini (2005) and the work on intertemporal price discrimination by a durable goods monopolist, such as Conlisk et al. (1984), Deb (2014), Garrett (2016) and Araman and Fayad (2020). The most closely related work is Garrett (2016) who studies a durable goods monopolist choosing a price path for dynamically arriving agents with changing values for a product and finds that cyclical pricing is optimal. A fundamental difference between our papers is the limited penalties the regulator in this paper has at her disposal.¹³ This constraint makes

¹²In the environmental hazard setting, the authors also show that it is without loss of generality to study mechanisms that induce immediate reporting and repair of the hazard. In the setting of this paper, the analogue of this is true only for low type agents.

¹³The regulator cannot punish above a maximum level, and faces a bound on the self-reporting incentive

the regulator’s problem in this paper non-trivial but precludes the use of techniques applied in Garrett (2016). Solving the model in this paper therefore requires a different approach that explicitly incorporates these limited penalties. Preferences in this paper also differ from those of the price discrimination setting, and this leads to different intuition underlying the optimal policy.¹⁴

This paper is not the first to note a link between amnesty and intertemporal price discrimination. Marchese and Cassone (2000) applies intuition and techniques from the literature on inter-temporal price discrimination to study a model of tax amnesty. Since values there are static, results are more closely connected to a monopolist discriminating between consumers with different but static values e.g. Conlisk et al. (1984).

In Section 6, I study a version of the model in which multiple agents commit crime together. This portion of the paper is closely related to the literature on leniency in multi-agent organizations, which has identified asymmetries in punishment as a force that can amplify the effects of self-reporting programs (see, for instance, Basu et al. (2016) in the case of corruption and Spagnolo (2000) in the case of cartels). Marvão and Spagnolo (2018) provides a survey of research on leniency programs for cartels, which offer lenient treatment only to the first firm (or firms) that reports their involvement. The most closely related papers are Motchenkova (2004), Harrington (2013), and Gärtner (2014) which study some dynamic aspects of leniency programs. Gärtner (2014) studies an environment in which a cartel’s risk of detection varies stochastically over time and observes that under cartel-preferred equilibrium selection, the overall effect of (static) amnesty is dampened when the process for detection can make large jumps. This dampening effect is important to the analysis of this paper as well and creates room for dynamic design. As we will see, multiplicity of equilibria becomes a relevant issue when studying the multi-agent setting. This latter part of the paper is therefore also related to the literature on full and virtual implementation, such as Maskin (1999) and Abreu and Matsushima (1992).

2. The Model

A stream of criminal agents (he) must decide whether to continue to operate or apply for amnesty. A regulator (she) chooses and commits to a penalty policy that is relevant for an agent’s decision. Calendar time is continuous, $t \in \mathbb{R}_+$.

she can offer. A natural lower bound on this incentive is that the regulator can at most offer not to punish a self-reporting agent at all, but the model allows any arbitrary bound.

¹⁴The regulator is concerned only with *fast* self-reporting by the agents. In the durable good monopoly setting, it would be as if the monopolist cared only that buyers purchased quickly, but not about the price.

2.1 The Agents

I present below the details of the agents' environment.

Arrival and Flow gain. Agents arrive according to a time homogeneous Poisson point process $M = (M_t)_{t \in \mathbb{R}^+}$ with rate parameter normalized to 1. Each agent is endowed with a *flow gain* process that follows a standard two-state continuous-time Markov chains $x_t^{t_0}$, independent across agents, with state space $E = \{x_l, x_h\}$ such that $0 \leq x_l < x_h$. For simplicity, state x_l is absorbing and agents transition from state x_h to x_l at Poisson rate λ . Upon arrival, agents are initialized in state x_h .¹⁵ I index $x_t^{t_0}$ by time since arrival so that $x_0^{t_0}$ is the initial state of an agent arriving at t_0 .

Choice and Detection. An agent arriving at time t_0 chooses a $[0, \infty]$ -valued stopping time with respect to the filtration generated by $(x_t^{t_0})_{t \geq 0}$, denoted τ , to irreversibly stop. I will use the terms *stopping* and *reporting* interchangeably, so that if an agent stops at some time t , I will also say that the agent reports his crime at t . For any stopping time τ , the calendar time at which the agent stops is $t_0 + \tau$. Upon stopping at calendar time t , the agent pays a terminal penalty $R_t \in [\underline{R}, \overline{R}]$ and his flow gains stop accruing. An agent is randomly detected by the regulator at time $t_0 + \tau_\rho^{t_0}$, where $\tau_\rho^{t_0}$ is an exponentially distributed stopping time with rate parameter ρ . The stopping time $\tau_\rho^{t_0}$ is independent of $(x_t^{t_0})_{t \geq 0}$ and across agents. If the agent is detected, he pays the maximum penalty \overline{R} and his flow gains stop accruing.

Payoffs. Combining the primitives described above, an agent's expected payoff from stopping time τ when arriving at time t_0 in state x_0 under penalty policy $R = (R_t)_{t \geq 0}$ is

$$W(x_0, t_0, \tau; R) := \mathbb{E} \left[\overbrace{\int_0^{\tau_\rho^{t_0} \wedge \tau} e^{-rs} x_s^{t_0} ds}^{(a)} - e^{-r(\tau_\rho^{t_0} \wedge \tau)} \left(\overbrace{\mathbf{1}_{\tau_\rho < \tau} \overline{R}}^{(b)} + \overbrace{\mathbf{1}_{\tau_\rho \geq \tau} R_{\tau+t_0}}^{(c)} \right) \middle| x_0 = x \right]$$

where the expectation is taken with respect to the distribution of $\tau_\rho^{t_0}$ and x_t . Note that λ does not explicitly appear, but rather controls the evolution of x_t . The first term, (a), represents the accrued flow gain until the minimum of (i) the time the agent chooses to stop and (ii) the time the regulator detects the agent. The second term, (b), is the penalty the agent receives when he is exogenously detected, \overline{R} , before choosing to stop. The third term, (c), is the penalty the agent receives when stopping before he is detected by the regulator.

¹⁵As I detail in Section 5, this assumption can be relaxed to allow for a time-independent arrival distribution across states without affecting the results.

The independence of $\tau_\rho^{t_0}$ from τ and $x_t^{t_0}$ implies that

$$W(x, t_0, \tau; R) = \mathbb{E} \left[\int_0^\tau e^{-(r+\rho)t} x_t^{t_0} dt - (1 - e^{-(\rho+r)\tau}) \frac{\rho}{\rho+r} \bar{R} - e^{-(\rho+r)\tau} R_{\tau+t_0} \middle| x_0 = x \right]$$

Note that the agent effectively discounts at rate $\rho+r$, which I call the *effective discount rate*. The agent solves the problem,

$$W^*(x, t_0; R) := \sup_{\tau \geq 0} W(x, t_0, \tau; R) \tag{A}$$

If a policy τ achieves value $W^*(x, t_0; R)$ it is called an *optimal stopping time* for the agent who arrives at time t_0 in state x . Where no ambiguity is possible, I drop the t_0 superscript from $x_t^{t_0}$ and $\tau_\rho^{t_0}$. Similarly, when it is clear what policy R is being used, I will write $W^*(x, t_0, \tau)$ and $W^*(x, t_0)$ rather than $W^*(x, t_0, \tau; R)$ and $W^*(x, t_0; R)$.

2.2 The Regulator

The regulator commits at time 0 to a *penalty policy* and an *obedient recommendation policy* i.e. a pair $((R_t)_{t \geq 0}, (A_t)_{t \geq 0})$ indexed by calendar time. The first component, $(R_t)_{t \geq 0}$, is a measurable function from \mathbb{R}_+ to $[\underline{R}, \bar{R}]$ with $\bar{R} \geq 0$.¹⁶ The second component, $(A_t)_{t \geq 0}$, is such that

- A_t is an element of $\{0, 1\}^{\{x_l, x_h\}}$
- $A^x(t) := A_t(x)$ is measurable for each x and
- $\tau := \inf\{t - t_0 \mid t \geq t_0 \text{ and } A_t(x_t) = 1\}$ is an optimal stopping time for an agent arriving at time t_0 .

The set of such policies $(R, A) := ((R_t)_{t \geq 0}, (A_t)_{t \geq 0})$ is denoted \mathcal{M} . The stopping time $\tau := \inf\{t - t_0 \mid t \geq t_0 \text{ and } A_t(x_t) = 1\}$ is called the *stopping time induced by A* for an agent arriving at t_0 .¹⁷

Payoffs. To calculate the regulator's payoff, I must track the characteristics of the population of agents. In particular, I need to know (i) how many are active, i.e. have not already chosen to report or been detected and (ii) the distribution across states x_h and x_l for active agents. Given a penalty policy R and a realization of:

- the population process $(M_t)_{t \geq 0}$,

¹⁶Note that I've placed no restrictions on \underline{R} . \underline{R} can, for instance, be negative and represent a reward, as in the case of gun buybacks. Nevertheless, because R_t here is pure money burning from the perspective of the regulator, the most natural cases involve $\underline{R} \geq 0$.

¹⁷Observe that the regulator is restricted to deterministic penalty policies (which I discuss in Section 5). This implies that A_t does not affect the agents' values, but rather (i) serves as a useful accounting device and (ii) allows the regulator to break ties in her favor.

- the flow gain process for each agent, $(x_t^{t_0})_{t \geq 0}$,
- and detection time $\tau_\rho^{t_0}$,

the stopping times induced by the obedient recommendation policy A induce a pair of stochastic paths, $(N_t^h)_{t \geq 0}$ and $(N_t^l)_{t \geq 0}$, describing the number of agents in states x_h and x_l at each time t , respectively. Let the distribution of $(N_t^h, N_t^l)_{t \in \mathbb{R}^+}$ induced by these processes be denoted $\mathbb{M}(R, A)$. The regulator discounts the future at the same rate as the agent, r , and her payoff from a policy (R, A) is:

$$V(R, A) := -\mathbb{E}_{\mathbb{M}(R, A)} \left[\int_0^\infty e^{-rs} (\alpha_h N_s^h + \alpha_l N_s^l) ds \right]$$

where $\alpha_h \in \mathbb{R}_{>0}$ and $\alpha_l \in \mathbb{R}_{\geq 0}$. The regulator solves the problem,

$$V^* := \sup_{(R, A) \in \mathcal{M}} V(R, A) \tag{\mathcal{P}}$$

A policy that achieves V^* is called *optimal*.¹⁸

3. Model Analysis

In this section, I compare dynamic to static policies and characterize an optimal policy. This is broken up into four steps:

- **Static Policies.** I define static policies as policies in which the terms of self-reporting are constant over time, and demonstrate their basic properties.
- **Low State Screening.** I show how to transform a policy into one which always induces immediate reporting by agents in the low state, under some conditions. This simplifies the subsequent analysis.
- **The Value of Dynamic Policies.** I characterize the set of model parameters for which a static policy is sub-optimal.
- **Optimal Policy.** I describe the main result of the paper, a characterization of an optimal policy. The optimal policy described takes a cyclical form.

¹⁸An optimal policy in the class is not necessarily the optimal mechanism in a general mechanism design approach, in which the regulator elicits reports from agents about their arrival time and returns from crime, and tailors self-reporting policies to these reports. I focus on this restricted class of policies, time paths for self-reporting penalties that apply uniformly to all agents, to remain as close as possible to the types of policies implemented in practice.

3.1 Static Policies

A static policy is one which is constant over time; it induces no added inter-temporal reporting considerations for the agent.

Definition 1. A static policy (R, A) is such that $R_t = v$ for all t for some $v \in [\underline{R}, \overline{R}]$. A dynamic policy is any policy which is not a static policy.

Denote by R^v the penalty policy in which $R_t = v$ for all t . An agent's decision problem under such a policy takes a particularly simple form. Consider first an agent in state x_l . This agent's decision is effectively static: his flow gains from crime, x_t , have been absorbed in state x_l and the penalty policy is a constant, v , each period. This agent need only consider two policies to compute his value: *immediately stop* ($\tau = 0$) or *never stop* ($\tau = \infty$). Immediately stopping delivers the agent a payoff $-v$, since he pays the penalty v and stops receiving his flow gain x_l . Never stopping delivers the agent a *gain* of $\frac{x_l}{\rho+r}$ — each instant he receives x_l , but he discounts at rate r and is detected at rate ρ — and a *loss* of $\frac{\rho\overline{R}}{\rho+r}$ — he is eventually detected, but because of discounting he only loses a fraction $\frac{\rho}{\rho+r}$ of \overline{R} . Combining these, an agent in the low state prefers to immediately stop *if and only if* $\frac{x_l - \rho\overline{R}}{\rho+r} \leq -v$.

Consider instead an agent in state x_h . Since $R_t = v$ for each t , while the agent remains in state x_h , his decision problem is identical across time. The agent need only consider three stopping policies in order to compute $W^*(x_h, t)$. First, the agent can stop immediately for a value of $-v$. Second, the agent can *never stop* i.e. choose $\tau = \infty$. This policy yields, for any arrival time t_0 , a value

$$W(x_h, t_0, \tau) = \frac{x_l - \rho\overline{R}}{\rho+r} + \frac{x_h - x_l}{\rho+r+\lambda}$$

where the first term is the value of this policy if the state were x_l and the second term is the extra value this policy delivers the agent given that he is in state x_h . The third policy is $\tau = \inf\{t - t_0 | t \geq t_0 \text{ and } x_t = x_l\}$ i.e. to stop immediately upon entering the low state. This yields a *gain* of $\frac{x_h}{\rho+r+\lambda}$ but a *loss* of $\frac{\rho\overline{R}}{\rho+r+\lambda} + \frac{\lambda v}{\rho+r+\lambda}$: the first term is the loss from detection and the second term is the loss from self-reporting in the low state. The value of this policy, for any arrival time t_0 , is then

$$W(x_h, t_0, \tau) = \frac{x_h - \rho\overline{R}}{\rho+r+\lambda} - \frac{\lambda v}{\rho+r+\lambda}.$$

Finding an optimal stopping time for the agent in state x_h simply requires comparing these three values.

The simplicity of the agent's decision problem under a static policy allows for a straightforward proof of the following proposition, which states that the optimal static policy is one which offers the agent the minimum penalty for self-reporting at all times i.e. \underline{R} . Recall that

R^v is the penalty policy such that $R_t = v$ for all t .

Proposition 1. *An optimal static policy is $(R^{\underline{R}}, A)$ where $A_t(x) = A_{t+s}(x)$ for each $x \in \{x_h, x_l\}$ and $t, s \in \mathbb{R}_+$.*

The proof is given in Appendix B but I provide here the intuition. Suppose that an agent considers reporting at time t or $t+s$. Lowering v has two effects from the agent's perspective: self-reporting at t becomes more valuable and so does self-reporting at $t+s$. However, the value to the agent from self-reporting at t increases by more than his value to self-reporting at $t+s$, because the increase at $t+s$ is scaled down by the agent's effective discount rate. Decreasing v as much as possible, i.e. to \underline{R} , is therefore optimal.¹⁹

3.2 Low Type Screening

In this section, I provide a lemma that allows me to focus on a sub-class of policies in the search for an optimal policy. I show that the regulator can induce reporting by low types at all times $t \in \mathbb{R}_+$ at no cost to the incentives of the high type agent.

Definition 2. *Let \mathcal{L} be the set of policies $(R, A) \in \mathcal{M}$ such that*

- $A_t(x_l) = 1$ for all t and
- $W(x_l, t, 0) = W(x_l, t, \tau)$ where $\tau = \inf\{s - t | s \geq t \text{ and } A_s(x_h) = 1\}$ i.e. an agent arriving in state x_l at time t is indifferent between reporting immediately and waiting to report until time $\inf\{s - t | s \geq t \text{ and } A_s(x_h) = 1\}$.

The first requirement is that an agent in the low state, x_l , chooses to report immediately. The second requirement is that an agent in the low state is indifferent between immediately reporting and instead waiting until $A_s(x_h) = 1$ to report i.e. following the high type's strategy as prescribed by A .²⁰ The following lemma states that any policy, (R, A) , can be transformed without loss for the regulator into a policy in \mathcal{L} . The lemma requires $x_l \leq \rho\bar{R} - (\rho + r)\underline{R}$ i.e. x_l is not so large that an agent in state x_l is unwilling to self-report when the penalty policy is $R^{\underline{R}}$. As we will see in Theorem 1, if this condition fails, a static policy is optimal. This immediately implies that either a static policy is optimal or the search for an optimal policy can be restricted without loss of value for the regulator to \mathcal{L} . Recall that $\mathcal{L} \subseteq \mathcal{M}$ and that if $(R, A) \in \mathcal{M}$ then A is an obedient recommendation policy.

Lemma 1 (Low Type Screening Lemma). *Suppose $x_l \leq \rho\bar{R} - (\rho + r)\underline{R}$. Then, for any policy $(R, A) \in \mathcal{M}$, there is another policy $(\tilde{R}, \tilde{A}) \in \mathcal{L}$ s.t. $A_t(x_h) = \tilde{A}_t(x_h)$.*

¹⁹It is also critical for the result that v , the penalty for self-reporting, does not enter the regulator's objective function explicitly.

²⁰Recall that agents are initialized with probability 1 in state x_h . Nevertheless, it is convenient to state the definition of \mathcal{L} in terms of the value of an agent arriving in state x_l .

The replacement penalty policy, \tilde{R} , has $\tilde{R}_t = R_t$ for any t such that the high type reports. Otherwise, the replacement penalty policy has $\tilde{R}_t = -W(x_l, t, \inf\{s-t | s \geq t \text{ and } A_s(x_h) = 1\})$, i.e. the negative of the low type's value of waiting to report until the next time at which the high type reports. The high type still finds it optimal to report at t such that $A_t(x_h) = 1$, since the value to any stopping time that stops with positive probability at t such that $A_t(x_h) = 0$ was available to him under the original policy as well.

To see an example, consider the policy in the left panel of Figure 1. In this policy, the

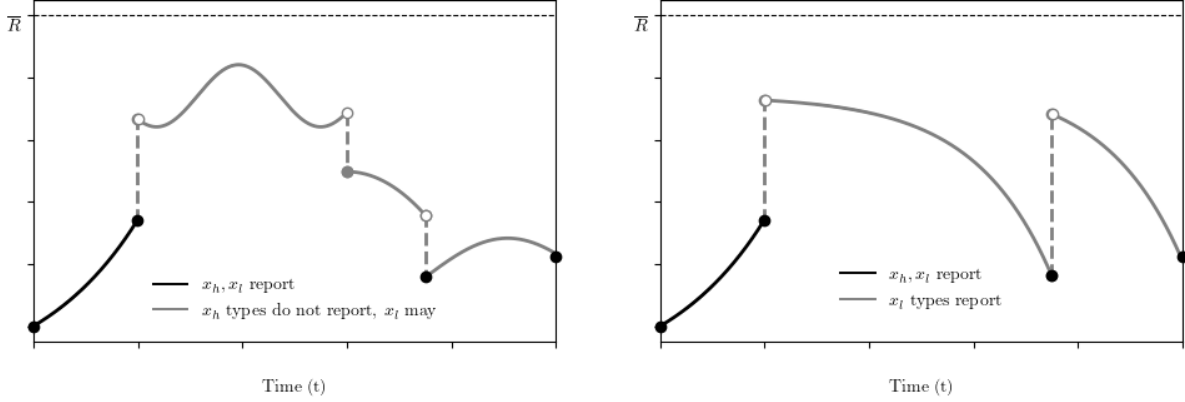


Figure 1: The left panel depicts the original policy. The right panel depicts the transformed policy, inducing the low type to report at all times without changing the reporting behavior of high types.

high and low types report on the black regions, while the low types may report elsewhere. There may also be times at which nobody reports. To clarify the transformation that the lemma performs, I describe a two-step process (the first step of which is not pictured) and abuse notation by using (R, A) for the original as well as the transformed policies. The first step of the transformation replaces all penalties on the gray region with \bar{R} and replaces $A_t(x_l) = 0$ at such points. This only strengthens the incentives of the agents to report on the black regions, since this step (weakly) worsens any alternative to immediate reporting. The result of the second (and last) step of the transformation is depicted in the right panel of Figure 1, where I replace the penalty at any time at which $A_t(x_h) = 0$ with $-W(x_l, t, \inf\{s-t | s \geq t \text{ and } A_s(x_h) = 1\})$, i.e. the negative of the low type's value from waiting until the next solid black point to report. I also set $A_t(x_l) = 1$ at such points. Reporting at any such point delivers the agent a value which was available to him in the original policy and so the new recommendation, which sets $A_t(x_h) = 1$ exactly on the black region and $A_t(x_l) = 1$ everywhere, is an obedient recommendation policy. The penalty policy in the right panel of Figure 1, along with the corresponding recommendation policy, is then a member of \mathcal{L} , with high types incentivized to report on the black region, as in the original

policy, and low types incentivized to report everywhere.

3.3 The Value of Dynamic Policies

With the previous lemma in place, I now answer the question: for what parameters can dynamic policies improve over static policies? The theorem below characterizes this set. Let $\theta := (\rho, r, \lambda, x_h, x_l, \bar{R}, \underline{R})$ denote an arbitrary collection of parameters and define Θ^* as follows:

Definition 3. Θ^* is the set of θ such that

$$V^* - \sup_{(R^v, A) \in \mathcal{M}} V(R^v, A) > 0$$

Θ^* is the set of parameters for which dynamic policies represent a strict improvement over static policies. I define now two objects which serve as boundaries in the characterization of Θ^* . First, let

$$\underline{\Theta}(\theta) = (\rho \bar{R} - (\rho + r)\underline{R} - x_l).$$

The term $\frac{\underline{\Theta}(\theta)}{\rho+r}$ is the value from *immediately reporting* ($-\underline{R}$) minus the value from *never reporting* ($\frac{x_l - \rho \bar{R}}{\rho+r}$), for an agent in state x_l when $R_t = \underline{R}$ for all t . As I argue below, when $x_h - x_l$ is larger than $\underline{\Theta}(\theta)$, a static policy will not induce agents in state x_h to report. Next, let

$$\bar{\Theta}(\theta) = \frac{\rho + r + \lambda}{\rho + r} \underline{\Theta}(\theta).$$

The pre-multiplying term $(\rho + r + \lambda)$, rescales the term $\frac{\underline{\Theta}(\theta)}{\rho+r}$ to make it comparable to the net returns from crime in the high state, $x_h - x_l$, which are killed not only by detection, ρ , and discounting, r but by transition to the low state λ . As I argue below, if $x_h - x_l$ is larger than $\bar{\Theta}(\theta)$, no policy can induce agents in state x_h to report.

Theorem 1. *The set of parameters θ for which dynamic policies strictly improve over static policies is non-empty and defined by the relation $\bar{\Theta}(\theta) \geq x_h - x_l > \underline{\Theta}(\theta)$, i.e.*

$$\Theta^* = \left\{ \theta \mid \bar{\Theta}(\theta) \geq x_h - x_l > \underline{\Theta}(\theta) \right\}$$

Observe that when $\lambda = 0$, i.e. high types do not transition to the low state, the set is empty, so dynamic policies do not improve over static policies.²¹ The result is proved in Appendix D. I provide here the intuition in the simple case when the regulator puts a 0

²¹This is not a consequence of the assumption that agents arrive in state x_h which, as I discuss in Section 5, can be generalized.

weight on low type agents in his value function (i.e. $\alpha_l = 0$), the gains in the low state are 0 (i.e. $x_l = 0$) and the minimum penalty the regulator can offer is 0 (i.e. $\underline{R} = 0$).

Suppose first that $x_h < \rho\bar{R}$ ($= \underline{\Theta}(\theta)$). Then $\frac{x_h}{\rho+r+\lambda} < \frac{\rho\bar{R}}{\rho+r+\lambda}$, which implies that the static policy $R^R = R^0$ already induces an agent in the high state to report. His total gain from waiting until the low state to report, $\frac{x_h}{\rho+r+\lambda}$, is not enough to offset his loss from detection under this strategy, $\frac{\rho\bar{R}}{\rho+r+\lambda}$. The high types report immediately upon arrival which is the first-best outcome for the regulator, and since the static policy achieves the regulator's first-best there is no need for dynamic policies.

Suppose instead that $x_h > \frac{(\rho+r+\lambda)}{\rho+r}\rho\bar{R}$ ($= \bar{\Theta}(\theta)$). Then, $\frac{x_h}{\rho+r+\lambda} > \frac{\rho\bar{R}}{\rho+r}$ which implies that x_h is so large that even if the agent were forced to make a once-and-for-all decision to immediately report for the minimum penalty possible, $\underline{R} = 0$, or continue until detected, he would choose to continue until detected. This implies that no policy can ever induce the agent to report in the high state; reporting yields a payoff $-\underline{R} = 0$, while continuing forever yields a strictly higher payoff.

Finally, suppose that $\bar{\Theta}(\theta) = (\rho+r+\lambda)(\frac{\rho\bar{R}}{\rho+r}) \geq x_h > \rho\bar{R}$ ($= \underline{\Theta}(\theta)$). In that case, under the static policy R^0 , an agent in the high state waits until he transitions to the low state to report. On the other hand, continuing forever is no longer an attractive option: if given a once-and-for-all opportunity, he would accept a reporting penalty of $\underline{R} = 0$ rather than continue forever. This once-and-for-all opportunity is defined as follows: $R_t = \bar{R}$ except at some $T > 0$, where $R_T = \underline{R} = 0$. This policy induces agents in the high state to report at T and strictly improves the regulator's value: a positive mass of agents in the high state chooses to report. The low types will no longer report everywhere as in the static policy, but this doesn't affect the regulator's value because $\alpha_l = 0$.

Allowing for $x_l > 0$ and $\underline{R} \neq 0$ requires only slightly different algebra. Allowing for $\alpha_l > 0$ requires application of Lemma 1, but is otherwise the same.

3.4 An Optimal Policy

The theorem that follows shows that a cyclical policy is optimal. Let T^* be the unique strictly positive solution in $T \in [0, \infty]$ to the equation

$$(x_h - x_l) \frac{1 - e^{-(r+\rho+\lambda)T}}{r + \rho + \lambda} - \frac{\rho\bar{R} - x_l}{\rho + r} (1 - e^{-(\rho+r)T}) - e^{-(\rho+r)T} \underline{R} = -\underline{R} \quad (\text{I})$$

if it exists and otherwise set $T^* = 0$. In the theorem below, T^* controls the frequency of cycles in the optimal policy described. As long as $\theta \in \Theta^*$, T^* is strictly positive. To help

understand the equation, it is written below in case $x_l = 0$:

$$\overbrace{x_h \frac{1 - e^{-(r+\rho+\lambda)T}}{r + \rho + \lambda}}^{\text{accumulated gains}} - \overbrace{\frac{\rho \bar{R}}{\rho + r} (1 - e^{-(\rho+r)T})}^{\text{potential losses from detection}} - \overbrace{e^{-(\rho+r)T} \underline{R}}^{\text{delayed reporting penalty}} = \overbrace{-\underline{R}}^{\text{immediate reporting penalty}}$$

This equation is a one-shot deviation constraint and guarantees that an agent in the high state is exactly indifferent between reporting at some time, s , with $R_s = \underline{R}$, and reporting at $s + T$ with $R_{s+T} = \underline{R}$. The first term in the equation on the left-hand side is the agent's accumulated expected gains from choosing to continue for a length of time T and the second term is the accumulated expected loss from being detected during the length of time T . The third term is the penalty that the agent will receive upon reporting for penalty \underline{R} at $s + T$, discounted to time s . The right-hand side is the payoff to immediate reporting at s .

The left-hand side of Equation (I) is displayed in Figure 2. The solid black line shows a parameterization in which $\theta \in \Theta^*$. The other two lines show two different parameterizations when $\theta \notin \Theta^*$, one in which x_h is very large and another in which it is very small.

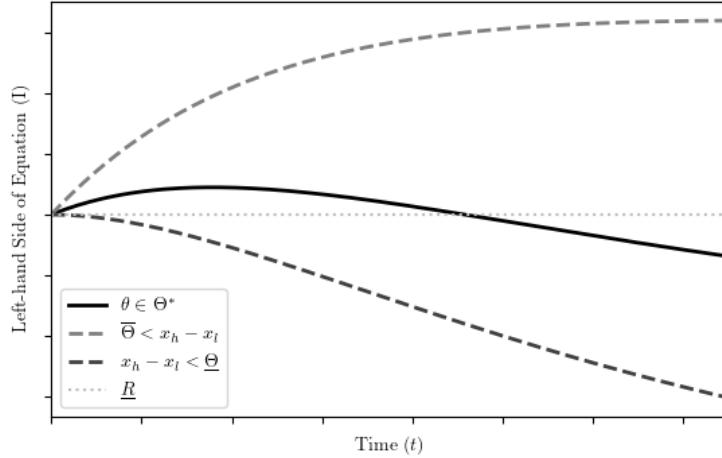


Figure 2: Left-hand side of Equation (I) when $x_l = 0$, for different parameters

In case $\theta \notin \Theta^*$, an optimal policy is $R_t = \underline{R}$ for all t . This case can happen in two ways. First, when $x_h - x_l < \underline{\Theta}(\theta)$, this policy induces reporting at all times by both high and low type agents. Second, when $x_h - x_l > \bar{\Theta}(\theta)$, this policy induces no reporting by high type agents, but no policy can. When instead $\theta \in \Theta^*$, the optimal policy takes a cyclical form. This is stated in the following theorem, along with details of the cyclical policy.

Theorem 2. *If $\theta \in \Theta^*$, an optimal policy, (R^*, A^*) , is:*

- $R_{nT^*+T_0}^* = \underline{R}$ for $n \in \mathbb{N}$ for some $T_0 \geq 0$
- $A_t^*(x_h) = 1$ if and only if $t \in \{T_0, nT^*\}_{n \in \mathbb{N}}$

- $R_t^* = e^{-(\rho+r)(T_0-t)} \underline{R} + (1 - e^{-(\rho+r)(T_0-t)}) \frac{(\rho \bar{R} - x_l)}{\rho+r}$ for $t < T_0$
- $R_t^* = e^{-(\rho+r)(T^* \lceil \frac{t}{T^*} \rceil - t)} \underline{R} + (1 - e^{-(\rho+r)(T^* \lceil \frac{t}{T^*} \rceil - t)}) \frac{(\rho \bar{R} - x_l)}{\rho+r}$ for $t > T_0$ and $t \notin \{T_0 + nT^*\}_{n \in \mathbb{N}}$
- $A_t^*(x_l) = 1$ for all $t \geq 0$

If $\theta \notin \Theta^*$, then an optimal penalty policy R^* is $R_t^* = \underline{R}$ for all t . An optimal obedient recommendation is constant i.e. $A_t(x) = A_s(x)$ for each $t, s \geq 0$ and $x \in \{x_h, x_l\}$.

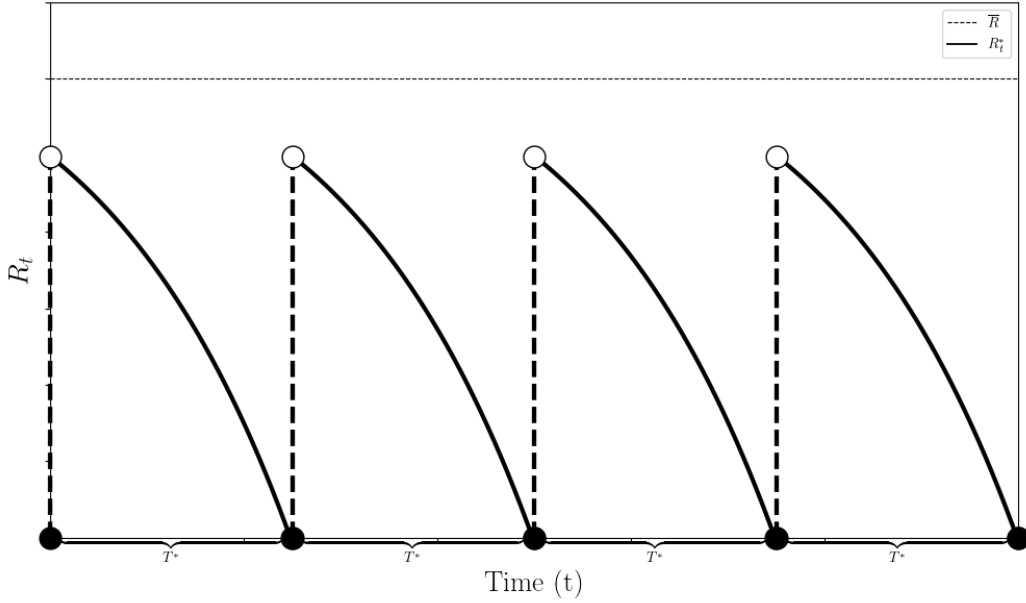


Figure 3: An optimal policy.

For $\theta \in \Theta^*$ and any $t \notin \{T_0, nT^*\}_{n \in \mathbb{N}}$, the path of R_t guarantees that an agent in state x_l is indifferent between immediately reporting anywhere on this path and waiting until the next $t \in \{T_0, nT^*\}_{n \in \mathbb{N}}$ to report. The T_0 in the theorem is an initial timing choice of the regulator, who is initially unburdened by incentives of prior agents. The optimal policy beyond T_0 is displayed in Figure 3. The economic intuition behind the use of such cyclical policies is that the agents *effectively discount* at rate $\rho + r$, the risk of detection plus time discounting, while the regulator discounts only at rate r . This creates a *backloading* motive on the part of the regulator. At any time t at which $A_t(x_h) = 1$, the regulator would like to incentivize agents to report by taking advantage of the discrepancy in discounting: she pushes future self-reporting rewards as far into the future as possible. The extent to which this can be done is restricted by the bounds on R_t , which leads, along with the result from Lemma 1, to the cyclical policy described in the theorem.

The proof of the theorem is given in Appendix E, but I detail here the steps and intuition. I first approximate the regulator’s problem by restricting her to obedient recommendation policies A such that for no pair t, s with $|t - s| < \epsilon$ is it true that $A_t(x_h) = A_s(x_h) = 1$. This allows the problem to be studied as if it were a discrete-time problem. As ϵ shrinks to 0, the value in this problem converges to the regulator’s value when $\epsilon = 0$. It turns out that this problem admits a policy that is optimal for any ϵ close enough to 0 (and hence for $\epsilon = 0$).

In this approximate problem, I solve first for the case in which $\alpha_l = x_l = 0$. This focuses the regulator’s problem on inducing high type agents to report. I then show that by applying Lemma 1, the general problem with $\alpha_l, x_l \geq 0$ can be transformed into a setting with new payoffs $x_l = 0$ and $\tilde{x}_h = x_h - x_l$, to which we can apply the proof for the case $\alpha_l = x_l = 0$.²²

To solve the $\alpha_l = x_l = 0$ case, I first restrict without loss of generality to policies such that the high type agent reports at t if and only if $R_t < \bar{R}$. This is without loss of generality because, at any point t at which the high type does not find it optimal to report, the regulator can replace R_t at that point with \bar{R} . This replacement only strengthens the incentives of the high type agents to report where they originally were. Although this may cause some low type agents not to report where they otherwise might have, the regulator does not lose value because $\alpha_l = 0$.

I show that a one-shot incentive compatibility condition can be derived as a generalization of Equation (I) – a high type agent should prefer to immediately report at t with $R_t < \bar{R}$ rather than wait until the next time s such that $R_s < \bar{R}$. For a given policy R , let $(T_i, R_{T_i})_{i \in \mathbb{N}}$ be the sequence of times and penalties such that $R_{T_i} < \bar{R}$. The high type agent finds it optimal to report at T_i for all i if and only if

$$P_k(T_{i+1} - T_i, R_{T_{i+1}}) \leq -R_{T_i} \quad \text{for all } i,$$

where

$$P_k(T, R) = x_h \frac{1 - e^{-(\rho+r+\lambda)T}}{\rho + r + \lambda} - \frac{\rho}{\rho + r} \bar{R} (1 - e^{-(\rho+r)T}) - e^{-(\rho+r)T} R.$$

The function $P_k(T, R)$ is a generalization of the left-hand side of Equation (I), allowing for penalties other than \bar{R} at the end of the waiting period T .

With this representation of the agent’s incentive compatibility condition, I reformulate the regulator’s problem recursively in the spirit of Spear and Srivastava (1987), using the promised penalty as a state variable. If at time T_{i+1} the state is $R_{T_{i+1}}$, the regulator is

²²The optimal policy I find here closely resembles the optimal policy in Garrett (2016), but requires different tools. In particular, the proof of the optimal policy in Garrett (2016) makes critical use of the fact that any sequence of sales dates (here, a sequence of reporting times for high types) can be induced by some price path, as well as the characteristics of the optimal price path that implements a given sales policy. In the setting of this paper, this relationship no longer holds as penalty paths that induce too frequent reporting by the high type constrain the set of strategies the regulator can implement in the future.

constrained to keep her promise to deliver the agent a penalty $R_{T_{i+1}}$ and a continuation that induces the agent to immediately report. Her choice at T_{i+1} is the next penalty, $R_{T_{i+2}}$, and the time until it is offered, $T_{i+2} - T_{i+1}$.

Abusing notation, let R denote an arbitrary state. I guess and then verify a solution to the recursive equation along with an optimal policy $(R'(R), T(R))$ that satisfies the constraints, where $T(R)$ is the time until high types are recommended to report again and $R'(R)$ is the penalty at that time. The guess that I propose is defined by the properties

- $R'(R) = \underline{R}$ and
- $T(R)$ is the strictly positive solution to $P_k(T(R), \underline{R}) = -R$.

Intuitively, this guess says that, whatever the promise the regulator must keep today, R , she should *always* keep it by offering \underline{R} as the next penalty, however long it takes to satisfy the incentive constraint. When $\theta \in \Theta^*$, the equation $P_k(T(R), \underline{R}) = -R$ has a unique strictly positive solution, so that $T(R)$ is well-defined.

To see why this guess is optimal, it is instructive to consider the sequential version of the regulator's problem with an initial state R . Between any two times T_i, T_{i+1} at which the high type reports, the regulator's loss is (Lemma E.1)

$$\int_0^{T_{i+1}-T_i} \alpha_h e^{-rt} \frac{1 - e^{-(\rho+\lambda)t}}{\rho + \lambda} dt$$

so this problem is

$$\begin{aligned} & \sup_{(R_{T_i}, T_i)_{i \in \mathbb{N}}} \sum_{i=0}^{\infty} \int_0^{T_{i+1}-T_i} \alpha_h e^{-rt} \frac{1 - e^{-(\rho+\lambda)t}}{\rho + \lambda} dt \\ \text{s.t. } & P_k(T, R_{T_{i+1}}) \leq -R_{T_i} \quad \forall i && \text{(Promise-keeping)} \\ & T_{i+1} \geq T_i + \epsilon, \\ & R_0 = R, \quad T_0 = 0, \quad R_{T_i} \in [\underline{R}, \bar{R}] \end{aligned}$$

After computing the integral and rearranging the sum, the problem becomes

$$\begin{aligned} & \sup_{(R_{T_i}, T_i)_{i \in \mathbb{N}}} C_2 + C_1 \sum_{i=0}^{\infty} x_h \frac{1 - e^{-(\rho+r+\lambda)(T_{i+1}-T_i)}}{\rho + r + \lambda} e^{-rT_i} \\ \text{s.t. } & P_k(T, R_{T_{i+1}}) \leq -R_{T_i} \quad \forall i && \text{(Promise-keeping)} \\ & T_{i+1} \geq T_i + \epsilon \\ & R_0 = R, \quad T_0 = 0, \quad R_{T_i} \in [\underline{R}, \bar{R}] \end{aligned}$$

for some C_1, C_2 with $0 < C_1$. Now relax the problem by dropping the constraints on R_t and

recursively substituting for R_{T_i} in the promise-keeping constraint to get,

$$\begin{aligned} & \sup_{(R_{T_i}, T_i)_{i \in \mathbb{N}}} C_2 + C_1 \sum_{i=0}^{\infty} x_h \frac{1 - e^{-(\rho+r+\lambda)(T_{i+1}-T_i)}}{\rho + r + \lambda} e^{-rT_i} \\ \text{s.t. } & \sum_{i=0}^{\infty} x_h \frac{1 - e^{-(\rho+r+\lambda)(T_{i+1}-T_i)}}{\rho + r + \lambda} e^{-(\rho+r)T_i} \leq C_3 - R \\ & T_{i+1} \geq T_i + \epsilon \end{aligned}$$

for some C_3 . In this problem, the i^{th} element of the sum in the objective is a *linear function* of the i^{th} element of the sum on the left-hand side of the constraint, discounted at a smaller rate: the agent discounts at his effective discount rate, $\rho+r$, while the regulator discounts at rate r . Intuitively, this suggests that the ideal way to satisfy the constraint is by *backloading* self-reporting incentives i.e. by setting T_0 very large, and setting $T_i = T_{i-1} + \epsilon$ for $i \geq 1$. In the correctly constrained problem, such a policy cannot satisfy promise-keeping constraints. Nevertheless, the guess for the optimal policy is a version of this constrained by the bounds on R_t . The cyclical path described in the theorem is the result of this backloading incentive meeting the bounds on R_t (along with the low type screening from Lemma 1).

Some properties of the guessed optimal policy are proved in Lemmas E.6 to E.8. The guess and associated value function are verified as optimal in Lemma E.9. I show that the regulator's and agents' value of a one-shot deviation from the proposed optimal amnesty policy can be represented by a function of *the effective discount rate* (time discounting for the regulator, and time discounting plus the risk of detection for the agent) that obeys a single-crossing property: if the function is positive for some discount rate, then it is strictly positive for any larger discount rate. This implies that if the value of a deviation is positive for the regulator (small discount rate) then it is strictly positive for the agent (high discount rate) and therefore violates the promise-keeping constraint, which held at equality under the guessed policy. The conclusion follows.

4. Applications

I have shown that optimal policies exhibit cycles of amnesty. I first discuss an applications of the model to military desertion and detail a case of desertion amnesties during the Russian Civil War. Afterwards, I discuss the model's implications for tax amnesties and gun buybacks and amnesties.

4.1 Desertion

From June 1919-June 1920 alone, the Red Army's *Central Anti-Desertion Commission* recorded over 2.6 million deserters, nearly equal to the number of new recruits over the same period (Figes, 1990). During the Vietnam War, over 400,000 soldiers deserted. The war minister of Napoleonic Italy declared desertion as "the first and principal obstacle to the organization of the army of the Kingdom" (Grab, 1995).

Desertion amnesties are often offered during the course of war in an effort to entice return and have been applied extensively across history.²³ The Red Army created its anti-desertion commission in 1918 – it increased punishments, strengthened enforcement (for instance, dispatching armed groups to search for deserters) and implemented periodic amnesties to entice deserters back to their units (Wright, 2012). As noted in Figes (1990), "...the most successful means of combating desertion [in the Red Army] were the amnesty weeks." During and surrounding the Argentine War of Independence, the military engaged in "alternating carrot and stick", offering amnesties to deserters in December 1813, September 1815, and September 1821 (Slatta, 1980). In Napoleonic Italy, "the government's repressive policy was mitigated by frequent amnesties designed to entice deserters and draft dodgers back to the army" (Grab, 1995). French Militaries in the 17th and 18th centuries offered periodic amnesties "interspersed with periods of severe repression, in an attempt to lure waverers back to their units" (Forrest et al., 1989).

It is not difficult to imagine how the difficulties that a deserter faces can change over time. Forrest et al. (1989) provides, among other things, an account of desertion in France in the early 19th century. Deserters were often hungry unless they were lucky enough to receive help from local people. Snow could block passages through the mountains and the cold could be deadly. Deserters were "forced into the surrounding countryside [of their village], searching out caves and hiding-places that would offer protection until the forces of law and order had passed through." Under such circumstance, Forrest et al. (1989) notes, "[I]t is hardly surprising that considerable numbers of deserters changed their minds."

Wright (2012) offers an account of the anti-desertion effort in the Red Army during the years 1918-1920, along with a detailed case study of the anti-desertion experience in Karelia. As noted in the case study, historians have deemed material shortages – 'uniforms, linen, tea, tobacco, and soap' – a primary reason for mass desertion during the Russian Civil War. Other important factors were the intensity of fighting, proximity to the White Army, and seasonality.²⁴ In response to the mass desertion problem, the Red Army created the Central Anti-Desertion Commission in December 1918. In June 1919, after an organizational period, the military introduced the use of periodic amnesties. Newspaper ads stressed the time-

²³In contrast to desertion amnesties offered after a war, as a method of reconciliation and forgiveness.

²⁴For instance, the harvest season led soldiers to return home to sow their fields.

limited nature of the amnesty, in order to encourage deserters' return.²⁵ During the months June to October 1919, multiple time-limited amnesty periods were offered, alongside harsh repression.

The model provides one lens through which to view the use of repeated, time-limited amnesties — by repeatedly offering an amnesty for only a short window, the anti-desertion program balanced two issues: (i) that deserters would not report unless their conditions became unbearable and (ii) that offering the program just once would ignore many deserters who would eventually be willing to re-enter the ranks. The application of amnesties, in this form, appeared to be a deliberate choice, rather than indecision — as noted in Rendle (2014), “One contemporary later argued that amnesties could have a significant impact as long as they were introduced when deserters were receptive, were not too frequent, and were applied alongside repression.”

It is instructive to consider the amnesties in the broader context in which they were offered. Aside from the general idiosyncratic variation in a deserter's plight as described earlier, some of the most important time-varying factors were the advances of the White army and the harvest season. In particular, deserters from the Red Army often returned at the end of their harvests, which is responsible for the success of some amnesties (Wright, 2012). A theory based on a public end to the harvest season would be able to account for annual amnesties, but even at a relatively small regional level, amnesties were more frequent — Wright (2012) describes a number of amnesty weeks separated by periods with no amnesty during the June-October 1919 period in the Karelia region. When instead the harvest timing is idiosyncratic (based on crop, sub-region, etc...), then the harvest season is less predictable and the theory presented in this paper applies.

Undoubtedly, a complete analysis of the amnesty-granting decision requires understanding the relationship between society, the military, and its personnel. As discussed in Wright (2012), the Red Army's overall decision to apply amnesty can be seen in part as a way of striking a balance between repression and restraint, in a bid to win the support of the peasantry. Nevertheless, this paper develops a formal model with a force, echoed in the writings of contemporaries, that drives towards the intermittent application of amnesties often observed as a response to the uncertainty and variation in the life of a deserter.

²⁵The following is snippet of a newspaper publication described in Wright (2012): ‘Deserters, townspeople! Today is the last day to appear before the commission [anti-desertion]. Hurry; present yourself today as tomorrow will be too late!’

4.2 Gun Amnesties and Buybacks

A typical gun amnesty program commits to a ‘no-questions’ asked acceptance of illegally owned firearms, freeing participants from the risks of illegal gun ownership.²⁶ Buy-back programs go one step further, offering to pay for each firearm surrendered. During the Argentinian buyback of 2007-08, the government collected more than 100,000 weapons (Lenis et al., 2010). During the Brazilian buy-backs of 2003, 2009 and 2011, the government collected more than 1 million weapons. When operated on a small scale, the evidence, especially in the U.S., points to the lack of any effect of gun buybacks on gun violence (Plotkin, 1996). On a large scale, however, these programs can potentially be effective (Lenis et al. (2010), Macinko et al. (2007)), especially when coupled with changes to the enforcement environment.

The inter-temporal properties of these programs vary considerably. Brazil, for instance, has operated a temporary buyback program four times since 2003. Sweden has operated three temporary buyback programs since 1993. Mexico operates a permanent buyback program. Tasmania operates a permanent amnesty program and all of Australia will begin to do so after 2020. In many cases, short-term gun buybacks are operated when public support is strong (e.g. after a tragedy) or when private funding is available (Plotkin, 1996).

The model in this paper explores one reason why the intermittent nature of some programs can be an advantage and how one can improve the design of programs that are offered continuously, such as Mexico’s gun buyback program. When the *option value* of participating in a gun-amnesty or buy-back is a first-order concern, the optimal amnesty program has both a permanent and temporary component – in a stylized setting, I have shown that an optimal policy induces self-reporting by agents with low returns from gun ownership at all times, but induces self-reporting by agents with high returns from gun ownership only intermittently. When instead this option value is not first-order, a static program is optimal. When illegal guns returned in amnesties/buy-backs have come into the owner’s possession innocuously – for example, through inheritance – the option value of amnesty is irrelevant and disposing the gun as soon as possible is, to a first-order approximation, the owner’s only objective. On the other hand, if the value to owning an illegal gun is derived from its operation by the owner, for safety, recreational or criminal reasons, then the owner has a more complicated objective: he would like to take advantage of an amnesty or buy-back when the weapon is no longer useful to him, but not before.²⁷ By offering only a limited-time buy-back, the regulator can entice a gun owner to self-report faster than he would under a

²⁶The exact content of ‘no-questions’ asked varies from program to program.

²⁷However, it must be noted that this change in value cannot come from a malfunction in the gun itself. As shown in Mullin (2001), such a change in value will only lead gun owners to turn their gun in during a buy-back only to turn around and use the money to buy a new gun.

permanent buy-back program.

Whether the optimal policy takes a dynamic form depends on parameters of the environment. While some parameters like the detection rate and penalties can be estimated from available data, the speed at which people transition from high to low value gun ownership cannot be. Many gun amnesty and buyback programs are accompanied with anonymous surveys of participants (McGuire et al., 2011). One way to estimate these parameters is to add two questions which are often left out of these surveys. The first is “how long have you owned your gun?” Answers to this question provide information on λ , the rate of transition from high to low value gun ownership. This can then inform the frequency and form of the optimal policy. The second is “if you owned it during the last buyback, why didn’t you turn it in then?”. Answers to this question can provide direct evidence on the motives of the participants. For instance, some may have not known about the program, or may have been otherwise misinformed about its conditions. Failing to account for such delay would overstate the value of a dynamic policy.

4.3 Voluntary Disclosure and Tax Amnesty Programs

Ptolemy V implemented the first recorded tax amnesty, circa 200 BC. Modern tax authorities have repeatedly implemented tax amnesties and voluntary disclosure programs including in the United States, Germany, Italy, India, the Phillipines, and Spain. Since 1980, more than 40 U.S. states have implemented a tax amnesty and 20 have implemented three or more. The use of tax amnesty is controversial, despite its prevalence (Le Borgne and Baer, 2008). Indeed, OECD (2015) states that tax amnesty programs, which it defines as programs that offer a reduction of the original tax amount, are “unlikely to deliver benefits that exceed their cost” (one of the reasons for which may be fairness considerations, which I discuss in Section 5). On the other hand, voluntary disclosure programs which offer reductions of penalties and interest and protection from prosecution can provide substantial benefits (OECD, 2015). In the model, such a constraint is best implemented by imposing $\underline{R} > 0$, representing the negative long-run effects on compliance and morale of programs which are too generous to evaders.

Le Borgne and Baer (2008) discusses the two main motivations for implementing a tax amnesty or voluntary disclosure program: “The two primary reasons for introducing tax amnesties are (i) to raise revenue in the short-term, and/or (ii) to increase compliance (e.g., by encouraging taxpayers to declare and pay previously undeclared tax, file tax returns, or register to pay taxes, so as to increase revenue and horizontal equity in the medium term).” It is often argued, especially recently, that tax amnesties have been implemented with an eye

to (i).²⁸ The model I present in this paper shows that despite this focus, the intermittent nature of these programs is also valuable from the perspective of (ii), increasing long-term compliance, when the value to tax evasion changes over time.

A basic question relevant for examining tax amnesties and voluntarily disclosure programs is, why do people apply? In the model presented, the change in gains from tax evasion leads evaders to self-report. Although direct evidence regarding motivation is not widely available, one source of evidence on this question comes from Ritsema et al. (2003), who implemented a survey of participants in the 2003 Arkansas tax amnesty program. The authors find that income, ease of evasion and inability to pay were three important determinants in the decision to evade taxes. These are factors which vary over time and therefore the model applies to tax settings in which these are first-order factors in deciding to apply for amnesty or to voluntarily disclose evasion. As discussed in Boise (2006), excessively frequent tax amnesties risk creating incentives for a tax evader to “forego participation in the first amnesty offered if he believes that another amnesty will be available in the foreseeable future.”

As detailed in OECD (2015), there are many examples of both permanent and repeated, temporary, tax amnesties and disclosure programs. Within the literature on tax evasion, the use of repeated, temporary amnesties has been a subject of some theoretical investigation. Marchese and Cassone (2000) rationalizes repeated tax amnesties as the tax authority price discriminating between honest and dishonest taxpayers, under the assumption that tax amnesties are implemented in cycles. Although the model in the present paper abstracts from important features of tax evasion and collection (importantly, that the regulator does not care about collected penalties), it offers a new take on the relative value of permanent versus repeated, temporary programs, focusing on how such programs reinstate those who have already decided to evade i.e. (ii) in the taxonomy of Le Borgne and Baer (2008). In this context, when the value to evasion is persistent and changes over time, it is sub-optimal to offer a static program and a cyclical program can provide stronger incentives for agents to self-report.

Andreoni (1991), in a one-shot setting, argues that a permanent partial voluntary disclosure program can be valuable as it provides agents with insurance against negative income shocks. One of the purposes of that paper is to rationalize some permanent programs observed in the United States and Canada. From a design perspective, the intuition presented in this paper suggests that such a program can be further improved when agents' values for tax evasion are imperfectly persistent, by offering only occasional opportunities for voluntary disclosure.

²⁸See Luitel and Tosun (2014).

5. Interpretation and Extensions

In this section, I provide interpretations of the assumptions in the model. I also detail some extensions and alternative modeling choices, as well as the model's limitations.

Arrival Time. Amnesty programs are valuable when the enforcement environment is too weak to deter all crime. In light of this, arrival by the agent to the model can be interpreted in two ways. First, the time the criminals begins committing the crime is perfectly observed (as in desertion during war), but the crime is initiated in a state in which the returns are too high to be deterred and only upon arrival to the model do they reach a level at which the regulator can induce self-reporting. If returns are private, arrival time to the model is then naturally viewed as private information of the criminal. As an example, consider the case of military desertion. A military will know, within a few hours, that a soldier has deserted. Therefore, the military can condition on desertion time when designing the amnesty program. For instance, the military could say that anyone who returns within two days will not be labeled a deserter and will be punished only mildly. Such programs could entice deserters who quickly decided they made a mistake, but not those who made calculated decisions and chose to remain deserters for longer periods.²⁹ In the Russian Civil War, the Red Army's desertion amnesties did not restrict the application of amnesty by desertion date. In other desertion settings, there have been such restrictions, but they are rare and often extend years in the past (such as the Sri Lankan desertion amnesty of 2008, which extended to all who deserted after 2005). Second, in other cases, it may be more natural to suppose that the regulator observes neither the private return nor the actual time at which the crime begins. This is the natural interpretation in, for instance, the gun amnesty and buyback case. Both of these interpretations accord with the view that enforcement is too weak to deter crime *ex ante* and so must focus on detection *ex post* through self-reporting.

To more formally interpret the model in the first way, I (i) suppose the regulator observes arrival time and chooses a time path for amnesty with this arrival time as time 0 and (ii) include a third state, $x_{hh} > x_h$, in which the agent begins his crime. The state x_{hh} is set so high that the agent would never self-report, reflecting the weak enforcement environment. At some Poisson rate, γ , he transitions from x_{hh} to x_h . While the regulator is allowed to condition her amnesty path on the time at which the agent begins crime, she does not observe the time of transition from x_{hh} to x_h . A formal analysis of this version of the model is provided in Appendix A.4. I show that as long as $\gamma < \rho$, the characterization of the optimal policy in Theorem 2 is unchanged.

²⁹For those, enforcement is too weak and offering an immediate amnesty is like the military asking them not to desert, which it already does.

In other settings, the regulator may be unsure of the precise arrival time of an agent, but has some information. To model this, one can suppose the regulator faces a single agent who arrives at some exponentially distributed time from 0, with parameter γ . This version of the model is technically equivalent to the version above and so as long as $\gamma < \rho$ the characterization of the optimal policy in Theorem 2 is unchanged.

Entry Deterrence. The model ignores the decision to become a criminal i.e. the extensive margin – criminals arrive criminals, as already described in the *Arrival Time* discussion above. A version of the model that introduces a third, very high state in which the agent arrives and chooses whether to become a criminal leaves the results unchanged. A richer model would have multiple states in which an agent can arrive to the model and decide whether to become a criminal; the agents arriving in the highest states cannot be deterred from crime, while agents arriving in the intermediate and low states can be. As agents transition from the high states down into the intermediate and low states, they engage with self-reporting programs. In this model, the optimal policy will trade off *entry deterrence* with *ex-post detection* from self-reporting. When the main motivation is entry deterrence, self-reporting programs will be counter-productive. This can happen when the distribution of entering values is concentrated below the level deterred by shutting down self-reporting programs. When the main motivation is ex-post detection, then self-reporting programs like the ones studied here will be useful. This can happen when the distribution of entering values is concentrated above the level deterred by shutting down self-reporting programs.

When the time at which the crime is committed can be observed and conditioned on (as in the discussion above on *Arrival Time*), then one simple way to deal with this is for the regulator to commit to a large T_0 , i.e. a long initial time without any amnesty, after which she can implement the optimal amnesty policy that we have discussed. If instead the time at which the crime is committed is not used in the description of the optimal policy, either because it is unknown or is too difficult to implement, the entry deterrence issue is more complicated. I investigate further in Appendix A.3.

Exogenous Detection Penalty. The penalty the agent receives when exogenously detected is \bar{R} . In this setting, this is without loss generality because the regulator does not place any value on collected penalties for their own sake.

Deterministic Policies. The regulator’s problem restricts her to *deterministic policies*. In Appendix A.1, I provide an extension of the main result to a restricted class of Poisson random policies. Nevertheless, I do not rule the possibility that general random policies improve the regulator’s value.

Note that the restriction to deterministic R ensures, within this class, it is without loss of generality for the agent to ignore A in computing her value W . If instead R is random, then the realization of A could convey relevant information and should be included as a conditioning variable in the agent’s problem.

Initial Distribution of Values. Arriving agents are initialized in state x_h . This can be relaxed to allow for a time-independent distribution of arriving values across the high and low states with no change to the results, but some extra burden on notation. As we have seen, an optimal mechanism induces low value agents to report either always or never. Allowing for a time-independent distribution of arriving values then just scales the regulator’s problem by a constant factor, leaving the optimal policy unchanged (like α_h).

Uncontrollable Rate of Detection. In the model, the regulator does not control the rate of detection. In Wang et al. (2016), a self-reporting problem is studied in the context of environmental regulation in which the regulator controls the rate of detection. In a different setting and in the absence of dynamic values, the paper shows that manipulating the risk of detection can amplify the value to dynamically adjusted self-reporting penalties. The model presented in this paper then fulfills a role complementary to Wang et al. (2016) by showing how self-reporting programs should behave in the absence of inspection control but in the presence of dynamic values to crime. It addresses the design of self-reporting programs in empirical settings in which the rate of inspection is indeed uncontrollable, unlike the Environmental Protection Agency (EPA) audit setting considered in Wang et al. (2016). For instance, in the case of price-fixing cartels, much of detection comes from buyer complaints which the anti-trust authorities do not directly control Harrington (2005). In the case of the anti-desertion campaigns in the Red Army, enforcement was locally delegated but deserters could be caught anywhere or discovered by people other than those tasked with explicit enforcement (Wright, 2012). In general, at least some detection typically comes from third-party reporting which the regulator cannot directly control.

Absorbing Low State. The model does not allow for the possibility that agents in the low state transition back to the high state. This assumption is made for tractability. Lemma 1 generalizes to this case and so a policy close to Theorem 2 can be shown to be approximately optimal when transitions back to the high state are small (with the loss relative to optimality shrinking with the size of the transition).

Collected Penalties. The model abstracts from the value of collected penalties — that is, the regulator’s objective function does not include collected penalties, treating R_t and \bar{R} as pure money burning.

In the case of tax amnesty, where one of the main motivations is short-term revenue, this issue is salient.³⁰ One way to incorporate revenue considerations is to generalize the regulator’s objective function to be a weighted combination of the loss from tax evasion (x_t) and the gain from penalties. When the weight on collected penalties is equal to the weight on the loss from operation, the game between regulator and agents is zero-sum: in this case the regulator minimizes the agents’ value, which is achieved by never granting a penalty reduction i.e. $R_t = \bar{R}$.³¹ When instead the regulator places a *lower* weight on the gain from penalties, there is scope for self-reporting to benefit the regulator. When the penalty represents prison time, a non-positive weight on penalties is apt.³² When penalties are financial, a lower weight represents the cost of collecting penalties and proving guilt, which is administratively expensive (Franzoni et al., 1996). Although I have no formal results for this setting, the main force at work remains in tact.

In the case of desertion, one interpretation of collected penalties is prison time. Militaries have found ways of preserving manpower while still imposing punishment, such as random punishment (Becker (1968), Chen (2017)), postponing prison sentences until after a war, relegating deserters to the worst duties, organizing penal battalions, and others. Nevertheless, a natural variation of the model would introduce a *loss* from collecting penalties, since by imprisoning a deserter, the military loses out on manpower. In this case, the incentive to offer self-reporting programs is even greater, since they give the regulator a way to avoid punishment and preserve manpower. The fundamental force of the paper is therefore strengthened in this context. The inclusion of this force complicates the analysis and how it affects the results is left as an open question. A more complex model may study how the government uses self-reporting incentives as a tool to speed up reporting and save on enforcement costs in conjunction with other tools.

In other cases, it is more natural to ignore the gains from collected penalties. In cartels, for instance, the penalties may be viewed as pure transfers, while anti-competitive behavior represents a dead weight loss (Motta and Polo, 2003). Valuing penalties from illegal gun ownership may be accommodated in a way similar to tax evasion, with different weights on penalties and behavior. But, gun ownership generates externalities through misuse, theft and unregulated sale (Cook and Ludwig, 2006); these are not internalized by the gun owner in weak enforcement environments. In this case, the weight the regulator places on collected penalties is small. Gun ownership penalties also involve prison time, which should not be treated positively in the regulator’s objective function.

³⁰See (Le Borgne and Baer, 2008) for a discussion of this issue.

³¹This is true in the single-agent model, but not the multi-agent model, where reporting by one agent topples the entire organization. In the latter case, equal weights still leaves scope for self-reporting to be valuable.

³²Or even a negative weight, since incarceration is costly.

Quitting. Although one of the real-life features motivating the model is that certain aspects of crime are irreversible, it is still interesting to think about a case in which an agent is given an option to “quit” without self-reporting. This may be especially important in cases like gun amnesties and buybacks, where it seems especially easy to dispose of or hide an illegal gun when it is not in use (and remove the evidence of its existence). When quitting is not free (because it is still risky to dispose of an illegal gun or hide it in a home where it may be mishandled), then the forces we have studied continue to apply, except the regulator is limited in how high a penalty she can entice a criminal to accept. I discuss further the possibilities when quitting and concealment, i.e. reversible quitting, are free in Appendix A. In short, when quitting is free, there is no role for amnesty when $\underline{R} \geq 0$ because an agent always weakly prefers to quit rather than self-report. When instead $\underline{R} < 0$, i.e. a buyback is feasible, then the basic forces remain.

Ethics. In some contexts with weak enforcement, the regulator may still enjoy a high rate of compliance. One reason posited for this in the case of tax compliance is ethics. Alm and Torgler (2011) argue that it is puzzling that so few citizens cheat on their taxes, given the relatively small risk of detection and penalties upon detection. The paper argues that ethics partially explains this high rate of compliance. Citizens comply because they perceive the tax system as fair, and the government as an institution that upholds this fairness. Generous tax amnesties may erode this sense of fairness and lead citizens who feel they have been unfairly treated (because they paid their taxes on time) to cheat on taxes in the future. In this way, a tax amnesty could create lasting damage to compliance.

Similar ethical forces could arise in other contexts as well. The use of amnesty may be particularly damaging in situations when enforcement is weak but compliance is high. Amnesties as described in this paper may then best be used in situations where fairness considerations are less relevant, or when a regulator can determine a minimum penalty level for amnesty that will not erode compliance through this channel.

6. Multi-Agent Setting

I study now an extension to multi-agent organizations, such as price-setting cartels. In a multi-agent setting, the regulator’s policy affects the risk that one agent takes through the self-reporting of another. For instance, if Bob and Anne are in a cartel and Bob reports to avoid being pre-empted by Anne, then withholding amnesty from Anne dampens Bob’s reporting incentives. Despite this issue, I show that dynamic policies can be valuable for the regulator under some restrictions on the set of available regulatory policies described in Section 6.2.

I derive the results in this section for the case in which x_t takes two values, as in Section 2. However, in Appendix A.5, I show that the results can be readily generalized to x_t that takes more than two values. This is in sharp contrast to the single-agent setting, where it is important for tractability that x_t takes only two values. This is in part a result of a strong unraveling force that appears in the multi-agent setting.

6.1 Model

I detail below the model of two agent criminal organizations.

Agents. Criminal organizations consisting of two agents arrive according to an arbitrary Poisson point process $(M_t)_{t \geq 0}$.³³ The flow gain, x_t , is common to both agents and perfectly observed in a given period. The process governing x_t is the same as in Section 2 — x_t can take two values, $x_h > x_l$, the process transitions from x_h to x_l at Poisson rate λ , and state x_l is absorbing. Criminal organizations arrive to the model in state x_h .

Choice and Detection. The regulator detects the entire organization (rather than just a single agent) at time homogeneous rate ρ . The agents in the organization play a stopping time game. When either the organization is exogenously detected by the regulator or at least one of the agents chooses to stop, the flow gains from x_t stop for *both agents*. Stopping in this setting is interpreted as *reporting involvement in the organization* to the regulator, and so I use the terms *stop* and *report* interchangeably. If the criminal organization is exogenously detected, each agent pays a terminal penalty \bar{R} . If instead one of the agents chooses to stop, the agents make payments according to a policy that will be described in the regulator’s section below.

A *pure strategy* for an agent is a stopping, or reporting time, τ and a *mixed strategy*, m , is a distribution over stopping times.³⁴ Expand the domain of W , the agent’s value, to include m , the stopping time distribution of i ’s partner in the organization. That is, $W(x, t_0, \tau; m)$, is an agent’s value when his organization arrives at t_0 in state x , he plays the stopping time τ and his partner plays the mixed strategy m .

A strategy profile is a pair (m_i, m_{-i}) . To conserve notation, a pure strategy profile is denoted (τ_i, τ_{-i}) , representing a mixed strategy m_i that puts probability 1 on τ_i . The solution concept used is Nash equilibrium.

³³Note that this is a generalization of the time homogeneous Poisson point process required in Section 2, a result of the strong unraveling forces present in the multi-agent setting that are absent in the single-agent setting

³⁴Formally, I enlarge the probability space governing x_t to include independent copies of $([0, 1], \lambda_i)$, where λ_i is Lebesgue measure. Then a mixed strategy m is a measurable map such that $m(\cdot, r)$ is Lebesgue a.e. a stopping time for each $r \in [0, 1]$. This follows Touzi and Vieille (2002).

Assumption: To simplify this section, I assume that

1. $x_h - (\rho + \lambda)\bar{R} \neq 0$
2. $\underline{R} = x_l = 0$
3. The agent does not discount exogenously, i.e. $r = 0$.

These assumptions are made for simplicity of calculation and the results do not depend on them. Note that because $\rho > 0$ acts like discounting, the definition of the agent’s problem requires no modification to deal with the last assumption.

Regulator. The regulator commits to a time path $((R_t)_{t \geq 0}, (p_t)_{t \geq 0})$, where $p_t = (p_{1,t}, p_{2,t})$ with $p_{i,t} \in [0, 1]$.³⁵ If an agent reports while his partner in the organization does not, the agent receives a terminal penalty of R_t , while his partner receives the full penalty \bar{R} . If the agents report at the same time, the regulator assigns probability $p_{i,t}$ to agent i receiving R_t and agent $-i$ receiving \bar{R} and probability $1 - p_{1,t} - p_{2,t}$ to both agents receiving \bar{R} . It need not be the case that $p_{i,t} = \frac{1}{2}$.

6.2 Simple Policies

As in Gärtner (2014), when the regulator can set $p_{i,t}$ general, she would set $R_t = 0$ and $p_{1,t} = p_{2,t} = 0$ for all t i.e. a penalty \bar{R} when players arrive simultaneously and $R_t = 0$ otherwise. This perpetually provides the maximum incentive to pre-empt one’s partner. This policy is both anonymous, in that it doesn’t require using the names of agents, and static as it does not require any change in the policy over time.

While a number of cartel leniency applicants have beat out their co-conspirators by just a few hours, many cases have just one leniency applicant or a long separation between the first and second. Different application times are a result of private information about one’s willingness to apply for leniency, a private evaluation of the cartel’s prospects, private beliefs about the risks of detection, strategic uncertainty, awareness of leniency, and beliefs about one’s partner’s awareness of leniency. Ideally, one would fully model the private information that leads to such non-simultaneous application. Nevertheless, it instructive to consider, in the current setting of complete information, how well the regulator could do if she chose not to condition her policy on the simultaneity of applications. This can be accomplished by setting $p_{1,t} = p_{2,t} = \frac{1}{2}$. This assumption appears often in the literature on leniency in cartels. For instance, Harrington (2008) and Landeo and Spier (2020) make such an assumption.

³⁵As in the single-agent model of Section 2, a primary motivation for the restriction on the set of mechanisms is staying as close as possible to the class of mechanisms applied in real world settings, such as in anti-trust.

Definition 4. (R, p) is a simple policy if $p_{1,t} = p_{2,t} = \frac{1}{2}$.

Such policies can be interpreted as nature randomly selecting one player, i , and applying the regulator's policy as if i arrived before $-i$. In the next section, Section 6.3, I provide a benchmark result when the regulator is completely unconstrained in her use of p . In Section 6.4, I study the problem instead when the regulator must use simple policies and show how to use a cyclical policy to achieve the same performance as in the fully unconstrained case.

6.3 Results: Unconstrained p

First, I consider a case in which the regulator can use $p_{i,t}$ and R_t freely – that is, she can condition her policy on simultaneous arrival. Consider the policy:

- $R_t^* = 0$ for all t
- $p_{1,t}^* = p_{2,t}^* = 0$

This policy applies a penalty of 0 if an agent is the first to apply, but applies a penalty of \bar{R} if players arrive simultaneously.

For any criminal organization, let $\tau^l := \inf\{t|x_t \leq x_l\}$ and $\tau^h := \inf\{t|x_t \leq x_h\}$. The latter, τ^h , is just immediate stopping.

Proposition 2. *Suppose $x_h > (\rho + \lambda)\bar{R}$. Then for any (R, p) , then (τ^l, τ^l) is a pure strategy equilibrium. Suppose instead that $x_h < (\rho + \lambda)\bar{R}$. Then, under policy (R^*, p^*) , any equilibrium $m^* = (m_1^*, m_2^*)$ is such that*

$$\mathbb{P}_{m^*}(\min\{\tau_1^*, \tau_2^*\} \leq \tau^h) = 1.$$

The general statement with multiple values of x_t is given in Proposition 2. The first part of the proposition states that when $x_h > (\rho + \lambda)\bar{R}$, regardless of the p and R the regulator uses, there is always an equilibrium in which agents do not report until reaching state x_l . The second part states that if $x_h < (\rho + \lambda)\bar{R}$, the policy (R^*, p^*) guarantees that in *every* equilibrium, agents immediately report by the time they reach state x^h . These two statements together imply that under *agent-preferred equilibrium* selection, the policy (R^*, p^*) is optimal.

6.4 Results: Simple Policies

In this section, I restrict the regulator to simple policies i.e. $p_{i,t} = \frac{1}{2}$ for each i and t . As I show below, static simple policies cannot, in general, achieve the performance of the policy in Proposition 2.

Proposition 3. Fix any simple static policy with $R_t = v$ for all t . If $x_h > (\rho + \frac{\lambda}{2})\bar{R}$, then (τ^l, τ^l) is a pure strategy equilibrium.

This is a special case of the more general result established in Proposition A.3. This proposition states that if $x_h > (\rho + \frac{\lambda}{2})\bar{R}$, the transition from state x_h to state x_l is sufficiently slow that each agent prefers to wait until state x_l to report, given that the agent's partner plans to do the same.

A cyclical policy that offers $R_t = 0$ only intermittently can, however, achieve exactly the performance of the unconstrained policy in Section 6.3. This cyclical policy is also anonymous i.e. does not condition on the name of the reporting agent. In this sense, the regulator experiences no loss from the use of simple anonymous policies as long as she is willing to use dynamic policies.

Starting at $t = 0$, define the policy,

1. Let $R_t^* := \bar{R} - \epsilon$ for a length of time T
2. Let $R_t^* := \bar{R} - \epsilon - t$ for a length of time $\bar{R} - \epsilon$
3. Return to (1) and repeat

For arbitrary ϵ, T , denote this policy by $R^*(\epsilon, T)$. An example of this type of policy is displayed in Figure 4. The slope of the downward sloping region is completely arbitrary. It serves only as an impetus for unraveling and any fixed slope will do.

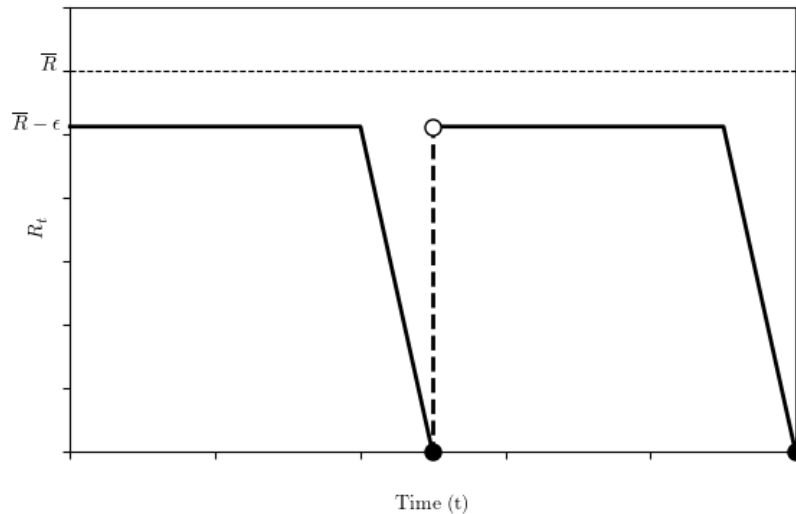


Figure 4: An Example of $R^*(T, \epsilon)$

Proposition 4. *Suppose $x_h < (\rho + \lambda)\bar{R}$. Then, there exists $R^*(\epsilon, T)$ such that if $m^* = (m_1^*, m_2^*)$ is an equilibrium, then*

$$\mathbb{P}_{m^*}(\min\{\tau_1, \tau_2\} \leq \tau^h) = 1.$$

This is a special case of Proposition A.4, presented in Appendix A.5. To see why the result holds, observe that at points $R_t = 0$, agents in state x_l find it strictly optimal to immediately stop, regardless of their partner's behavior. Since the same is true for the agent's partner, this leads to an unraveling that implies that any equilibrium that arrives with probability > 0 to state x_l must involve immediate reporting by the agents. Suppose then that $x_h < (\rho + \lambda)\bar{R}$. At a point $R_t = 0$, an agent knows that if he and his partner transition to state x_l during the $R_t = \bar{R} - \epsilon$ region, he and his partner will immediately report. This yields an expected penalty only slightly lower than \bar{R} , for sufficiently small ϵ . As long as T is large enough, this is nearly certain to happen. So, an agent reports at a point $R_t = 0$ when in state x_h . This leads to a complete unraveling through time, so that any equilibrium requires immediate reporting in state x^h with probability 1.

The previous two propositions lead to a simple corollary,

Corollary 1. *If $x_h - \rho\bar{R} \in (\frac{\lambda}{2}\bar{R}, \lambda\bar{R})$, then for any simple static policy (R, p) , (τ^l, τ^l) is an agent-preferred equilibrium and every agent-preferred equilibrium, (τ_1^*, τ_2^*) , is such that*

$$\mathbb{P}(\min\{\tau_1^*, \tau_2^*\} \leq \tau^l) = 1.$$

On the other hand, there exists $R^(T, \epsilon)$ such that in every equilibrium (τ_1^*, τ_2^*) ,*

$$\mathbb{P}(\min\{\tau_1^*, \tau_2^*\} \leq \tau^h) = 1.$$

To see this, observe that if there exists an equilibrium in which each agent uses a strategy that is almost surely more delayed than in any other equilibrium, then it is also an agent-preferred equilibrium. Combining this statement, the proposition above, and Proposition 3 delivers the corollary.

Finally, we can apply Proposition 2 to find that under agent-preferred equilibrium selection, the regulator's outcome under $R^*(T, \epsilon)$ is no more delayed than when she can use general, non-simple policies.

6.4.1 Anti-Trust

Cartels are multi-firm organizations that fix prices, allocate markets, or otherwise explicitly agree to restrain competition. One of the most powerful tools that the Department of Justice and other anti-trust authorities use in the fight against cartels is the leniency program. This program offers immunity from prosecution and fines to the first member of a cartel to

admit involvement to the anti-trust authority (and sometimes even protection from follow-on private litigation).

In all major jurisdictions, the leniency program is static; at any time, a member of a cartel can apply for amnesty and receive immunity from fines and prosecution.³⁶ However, demand and cost variation as well as entry can lead to swings in the value to operating a cartel. Furthermore, cartelists rarely report simultaneously and reporters are not treated differently depending on how soon before or after their partner they report. In the context of the model presented here, then, an optimal program may not be static. Proposition 4, and its generalized version Proposition A.4, shows that if the risk of market events that erode profits is sufficiently high, a static simple amnesty can be improved on by one which offers amnesty only intermittently. Whether such a program indeed offers an improvement depends on the specifics of the anti-trust environment, in particular the parameters of the process x_t . Nevertheless, this paper provides a first step in analyzing the role that dynamic policies can play when static policies do not lead to complete unraveling.

The class of policies considered uses no information about the cartel in determining the eligibility of a member for amnesty — not the time the cartel is born, nor anything about its profitability. The policy also does not require conditioning on the simultaneous arrival of the agents. Therefore, the policy prescribed in Proposition A.4 satisfies the need for transparency and predictability which regulators have described as a critical piece of a successful leniency program.³⁷

7. Discussion

I conclude now with a recap and discussion of the baseline model of Section 2 and the multi-agent model of Section 6.

7.1 Baseline Model

In Section 2, I study the problem of a regulator who designs amnesty programs to induce reporting by criminal agents. I show (Theorem 1) that when the returns from crime can change over time ($\lambda > 0$), the generosity of an optimal amnesty program may change over time as well. In such cases, Theorem 2 establishes that the optimal policy is cyclical and describes its form. Except for an initial period, the minimum possible penalty (\underline{R}) for

³⁶There are, however, other dynamic features of the leniency program. For instance, the penalty imposed on a cartelist increases with his estimated gains from cartelizing, up to a limit of 10% of the firm's annual turnover. Furthermore, in some jurisdictions, even cartelists who are not the first to report can receive some leniency, if they provide evidence important to the investigation.

³⁷<https://www.justice.gov/atr/speech/cornerstones-effective-leniency-program>

reporting is offered at evenly spaced points in time. In between such times, a decreasing schedule of penalties is offered. Agents with a high return from crime report only at the end of each cycle while those with a low return from crime report at all times during the cycle. Along with simplifying features of the model, a *backloading* motive on the part of the regulator drives the optimality of this form of amnesty. Agents effectively value the future less than the regulator, even when exogenous discounting is the same, because they may be detected (at rate ρ). The regulator therefore finds it optimal to push the rewards from self-reporting as far into the future as possible. The optimal form of amnesty is a result of this backloading motive meeting the bounds on the penalty.

There are a number of avenues for future work. First, it would be useful to study a version of the problem in which the regulator can control, at some cost, the rate of detection. Second, new insights might result from incorporating political economy constraints into the model. In particular, the regulator may not be able to fully commit to a policy because she is occasionally replaced by a new regulator, wiping away previous commitments (as in a case of tax amnesty where the government is replaced every few years). Third, it would be interesting to further examine the entry deterrence margin beyond what has been discussed in Section 5. For instance, when the regulator cannot condition her policy on the time at which crime begins, randomization can be useful — by randomizing the timing of amnesty, agents cannot take advantage by initiating crime at times close to attractive amnesties. In that case, how should the regulator randomize amnesty?

7.2 Multi-Agent Model

The baseline model of Section 2 is extended in Section 6 to a setting with multiple agents i.e. a *criminal organization*. Dynamic amnesty programs that change over time can be valuable for the regulator, taking advantage of the powerful unraveling force that strengthens static amnesty programs by inducing a preemption effect.³⁸ I show, in particular, that when the returns from crime in the organization can make downward jumps, a dynamic amnesty program can sometimes induce more reporting than a static amnesty program, with the gains from the dynamic program magnified through the preemption effect.³⁹ In essence, the static amnesty program generates a powerful unraveling force on which it does not fully capitalize. I show that, under some parameterizations and with two values for the returns from crime, a simple cyclical amnesty policy can increase the amount of unraveling in equilibrium. In Appendix A.5, I show that this can be extended to a return process that takes more than

³⁸For the large literature that studies this force in the context of cartels, see the survey in Marvão and Spagnolo (2018).

³⁹When the returns from crime are perfectly observed, common value and move continuously over time, agent-preferred equilibria involve either immediate reporting or no reporting. See for instance Gärtner (2014). Unraveling would be weaker in a model with private information or boundedly rational agents.

two values.

In practice, the majority of cartel members play this game once. The results of this section can be similarly stated with rationalizability as a solution concept. In this case, however, generating the strong unraveling force through time and states requires common knowledge of rationality. A model that relaxes such rationality assumptions, perhaps by using a model of k -level reasoning, is a possibly fruitful way to bring these results closer to application.

The multi-agent presented leaves a number of interesting questions left unanswered. What if agents, under a static policy, are not playing their most preferred equilibrium? If the static policy induces much more unraveling than in the agent preferred equilibrium, then disturbing this regime by introducing the policy of Proposition A.4 can worsen the regulator’s payoff. Are there alternative reasonable equilibrium selection mechanisms? Is it possible to develop empirical tests to measure how delay is affected by the use of different amnesty policies?

I have assumed, as in the single-agent model, that organizations cannot stop without reporting i.e. collectively *quit*. The cyclical policy $R^*(\epsilon, T)$ has the feature that at some point, the agents will receive full amnesty if they report. So although collectively quitting may minimize the risk of detection, the unraveling effect will remain as long as there is some chance that the cartel is detected even after it stops.⁴⁰

Beyond this, the assumptions of the multi-agent setting presented here are extreme — agents are symmetrically informed about a pure common value return. Extensions to allow for private information and private values are natural, especially in cases of vertical organizations or organizations in which individuals are otherwise differentiated. Future work would augment the model with both of these. I have also abstracted from the interaction between self-reporting programs and the need for collusive organizations to be self-sustaining. Modeling this interaction, as has been done in the context of static leniency programs in anti-trust, is an interesting avenue for future work. Finally, although the results are presented for a case in which x_t varies exogenously over time, related results could be established for a case in which ρ varies exogenously over time.

References

ABREU, D. AND H. MATSUSHIMA (1992): “Virtual Implementation in Iteratively Undominated Strategies: Complete Information,” *Econometrica: Journal of the Econometric*

⁴⁰Evidence from cartels discovered through the leniency program in Europe suggests that some leniency applicants report their involvement long after the collapse of the cartel. See Zhou and Gärtner (2012) for further details.

- Society*, 993–1008.
- ALM, J. AND B. TORGLER (2011): “Do ethics matter? Tax compliance and morality,” *Journal of Business Ethics*, 101, 635–651.
- ANDREONI, J. (1991): “The Desirability of a Permanent Tax Amnesty,” *Journal of Public Economics*, 45, 143–159.
- ARAMAN, V. F. AND B. FAYAD (2020): “Intertemporal price discrimination with time-varying valuations,” *Operations Research*.
- BASU, K., K. BASU, AND T. CORDELLA (2016): “Asymmetric punishment as an instrument of corruption control,” *Journal of public economic theory*, 18, 831–856.
- BATTAGLINI, M. (2005): “Long-term contracting with Markovian consumers,” *American Economic Review*, 95, 637–658.
- BECKER, G. S. (1968): “Crime and punishment: An economic approach,” in *The economic dimensions of crime*, Springer, 13–68.
- BOISE, C. M. (2006): “Breaking Open Offshore Piggybanks: Deferral and the Utility of Amnesty,” *Geo. Mason L. Rev.*, 14, 667.
- CHEN, D. L. (2017): “The deterrent effect of the death penalty? Evidence from British commutations during World War I,” *Evidence from British Commutations During World War I (Feb 26, 2017)*.
- CONLISK, J., E. GERSTNER, AND J. SOBEL (1984): “Cyclic Pricing by a Durable Goods Monopolist,” *The Quarterly Journal of Economics*, 99, 489–505.
- CONNOR, J. M. (2007): *Global price fixing*, Springer.
- COOK, P. J. AND J. LUDWIG (2006): “The Social Costs of Gun Ownership,” *Journal of Public Economics*, 90, 379–391.
- DEB, R. (2014): “Intertemporal Price Discrimination with Stochastic Values,” *Unpublished manuscript, University of Toronto*.
- DREYFUS, P., L. E. GUEDES, B. LESSING, A. R. BANDEIRA, M. DE SOUSA NASCIMENTO, AND P. S. RIVERO (2008): “Small Arms in Rio de Janeiro,” *The Guns, the Buypack, and the Victims*.
- FIGES, O. (1990): “The Red Army and Mass Mobilization during the Russian Civil War 1918-1920,” *Past & Present*, 168–211.

- FORREST, A. I., A. J. F. FORREST, ET AL. (1989): *Conscripts and Deserters: The Army and French Society During the Revolution and Empire*, Oxford University Press on Demand.
- FRANZONI, L. ET AL. (1996): “Punishment and Grace: on the Economics of Permanent Amnesties,” Tech. rep.
- GARRETT, D. F. (2016): “Intertemporal Price Discrimination: Dynamic Arrivals and Changing Values,” *American Economic Review*, 106, 3275–99.
- GÄRTNER, D. L. (2014): “Corporate Leniency in a Dynamic World: The Preemptive Push of an Uncertain Future,” *Available at SSRN 2340973*.
- GRAB, A. (1995): “State power, brigandage and rural resistance in Napoleonic Italy,” *European History Quarterly*, 25, 39–70.
- HARRINGTON, J. E. (2005): “Detecting cartels,” Tech. rep., working paper.
- (2008): “Optimal corporate leniency programs,” *The Journal of Industrial Economics*, 56, 215–246.
- (2013): “Corporate Leniency Programs when Firms have Private Information: The Push of Prosecution and the Pull of Pre-emption,” *The Journal of Industrial Economics*, 61, 1–27.
- KAPLOW, L. AND S. SHAVELL (1994): “Optimal Law Enforcement with Self-Reporting of Behavior,” *Journal of Political Economy*, 102, 583–606.
- KRASIKOV, I., R. LAMBA, AND T. METTRAL (2019): “Implications of unequal discounting in dynamic contracting,” *Available at SSRN 3142932*.
- LANDEO, C. M. AND K. E. SPIER (2020): “Optimal Law Enforcement with Ordered Leniency,” *The Journal of Law and Economics*, 63, 71–111.
- LE BORGNE, E. AND K. BAER (2008): *Tax Amnesties: Theory, Trends, and Some Alternatives*, International Monetary Fund.
- LEIGH, A. AND C. NEILL (2010): “Do Gun Buybacks Save Lives? Evidence from Panel Data,” *American Law and Economics Review*, 12, 509–557.
- LENIS, D., L. RONCONI, AND E. SCHARGRODSKY (2010): “The Effect of the Argentine Gun Buy-back Program on Crime and Violence,” in *Universidad Torcuato Di Tella. Argentina Retrieved from http://socrates.berkeley.edu/~raphael/IGERT/Workshop/PEVAF_September_27_2010.pdf*.

- LUITEL, H. S. AND R. S. SOBEL (2007): “The Revenue Impact of Repeated Tax Amnesties,” *Public budgeting & finance*, 27, 19–38.
- LUITEL, H. S. AND M. S. TOSUN (2014): “A reexamination of state fiscal health and amnesty enactment,” *International Tax and Public Finance*, 21, 874–893.
- MACINKO, J. ET AL. (2007): “Reducing Firearm Injury: Lessons from Brazil.” *LDI issue Brief*, 12, 1–4.
- MALIK, A. S. (1993): “Self-reporting and the Design of Policies for Regulating Stochastic Pollution,” *Journal of Environmental Economics and Management*, 24, 241–257.
- MARCHESE, C. AND A. CASSONE (2000): “Tax Amnesty as Price-Discriminating Behavior by a Monopolistic Government,” *European Journal of Law and Economics*, 9, 21–32.
- MARVÃO, C. AND G. SPAGNOLO (2018): “Cartels and Leniency: Taking Stock of What We Learnt,” in *Handbook of Game Theory and Industrial Organization, Volume II*, Edward Elgar Publishing.
- MASKIN, E. (1999): “Nash Equilibrium and Welfare Optimality,” *The Review of Economic Studies*, 66, 23–38.
- MCGUIRE, M., M. MANNO, A. ROOK, L. MARANDA, E. RENAUD, A. DEROSS, AND M. HIRSH (2011): “Goods for guns—the use of a gun buyback as an injury prevention/-community education tool,” *Journal of Trauma and Acute Care Surgery*, 71, S537–S540.
- MOTCHENKOVA, E. (2004): “Effects of Leniency Programs on Cartel Stability,” *CENTER discussion paper*.
- MOTTA, M. AND M. POLO (2003): “Leniency programs and cartel prosecution,” *International journal of industrial organization*, 21, 347–379.
- MULLIN, W. P. (2001): “Will Gun Buyback Programs Increase the Quantity of Guns?” *International Review of Law and Economics*, 21, 87–102.
- OECD (2015): “*Update on Voluntary Disclosure Programmes: A pathway to tax compliance*”.
- PLOTKIN, M. R. (1996): in *Under Fire: Gun Buy-Backs, Exchanges and Amnesty Programs*, Police Executive Research Forum Washington, DC.
- RENDLE, M. (2014): “Mercy Amid Terror? The Role of Amnesties During Russia’s Civil War,” *Slavonic & East European Review*, 92, 449–478.

- RITSEMA, C. M., D. W. THOMAS, AND G. D. FERRIER (2003): “Economic and Behavioral Determinants of Tax Compliance: Evidence from the 1997 Arkansas Tax Penalty Amnesty Program,” in *IRS research Conference*, Citeseer.
- SHIRYAEV, A. N. (2007): *Optimal Stopping Rules*, vol. 8, Springer Science & Business Media.
- SPAGNOLO, G. (2000): “Optimal leniency programs,” *FEEEM Working Paper*.
- SPEAR, S. E. AND S. SRIVASTAVA (1987): “On Repeated Moral Hazard with Discounting,” *The Review of Economic Studies*, 54, 599–617.
- TOUZI, N. AND N. VIEILLE (2002): “Continuous-time Dynkin games with mixed strategies,” *SIAM Journal on Control and Optimization*, 41, 1073–1088.
- WANG, S., P. SUN, AND F. DE VÉRICOURT (2016): “Inducing Environmental Disclosures: A Dynamic Mechanism Design Approach,” *Operations Research*, 64, 371–389.
- WILLMER, M. (1971): “On the Estimation of the Number of Persons Entering and Leaving Crime Annually,” *Journal of Research in Crime and Delinquency*, 8, 175–181.
- WOSEPKAT, J. L. (1971): “Repatriation and the Chieu Hoi Amnesty Approach in Vietnam: Consequences and Prospects,” *The International Lawyer*, 637–651.
- WRIGHT, A. S. (2012): “Stemming the Flow: The Red Army Anti-Desertion Campaign in Soviet Karelia (1919),” *Revolutionary Russia*, 25, 141–162.
- ZHOU, J. AND D. L. GÄRTNER (2012): “Delays in Leniency Application: Is There Really a Race to the Enforcer’s Door?” .

Appendix

A. Extensions

In this section I discuss a number of features that are missing in the model.

A.1 Poisson Randomization

Until now, I have restricted myself to the search for optimal deterministic policies. Although I don’t characterize the optimal policy for general random policies, I expand the model to

allow for a limited class of random policies and show that the deterministic optimal policy remains optimal. First, extend V linearly to random policies.

I prove the result for the case of $\underline{R} = 0$, but it extends readily to the case of general \underline{R} . Consider then the following class of policies:

Definition 5. A randomized policy (R, A) is called a γ -Poisson policy if there exists T_0 and a sequence of random variables $(T_i)_{i \geq 0}$ s.t.

- $T_{i+1} - T_i$ independent of T_i and exponentially distributed with rate parameter γ
- $R_{T_i} = 0$ for $i \in \mathbb{N}$ and $R_t = \bar{R}$ otherwise
- $A_t(x_h) = 1$ if and only if $t \in \{T_0, T_i\}_{i \in \mathbb{N}}$

The set of Poisson policies is denoted Γ .

These policies feature inter-arrival times of complete amnesty that are exponentially distributed with mean $\frac{1}{\gamma}$.⁴¹ I restrict to the setting in which $x_l = \alpha_1 = 0$ and argue that the policy in Theorem 2 remains optimal when allowing the regulator to choose from Γ . Let $\mathcal{M}^\Gamma := \mathcal{M} \cup \Gamma$.

$$V^\Gamma := \sup_{(R,A) \in \mathcal{M}^\Gamma} V(R, A) \tag{\mathcal{P}^\Gamma}$$

i.e. the expanded regulator’s problem allowing for policies in Γ (with some abuse of notation since R and A are now random variables).

Theorem A.1. Suppose $x_l = \alpha_1 = 0$. Then the optimal policy of (\mathcal{P}) defined in Theorem 2 is optimal in (\mathcal{P}^Γ) .

The proof is given in Online Appendix I. Random mechanisms in Γ which are incentive compatible for the agent require putting substantial probability on long periods without amnesty, relative to the deterministic mechanism with inter-arrival times T^* . As before, since the effective discount rate of the agent is higher than the discount rate of the regulator, this sampling is more costly for the regulator than the agent.

A.2 Quitting and Concealment

In many cases, such as desertion, it is difficult or impossible to “quit”. In cases such as tax amnesty, an evader can still be caught even years after the evasion took place (the limitations on this vary across jurisdictions) and, furthermore, adjusting tax returns to become compliant

⁴¹Notice that the choice of T_0 is as in problem (\mathcal{P}) to ensure that this initial choice doesn’t drive any differences between the two problems.

could raise suspicion and risk detection by tax authorities. In tax amnesties geared towards repatriation of legal foreign-held assets (as in the U.S. American Jobs Creation Act of 2004), repatriating assets outside of the amnesty period would be detected by tax authorities.

In other cases, such as gun amnesty, “quitting” illegal gun ownership can be accomplished by illegally selling the gun or throwing it in the river, which entails high penalties if caught in action, or keeping the gun hidden so that it goes undetected by authorities. Incorporate quitting into the model mutes the basic force that drives results, since agents now have a way to leave crime that does not involve self-reporting, but they still remain under certain conditions. There are a number of ways to model quitting, but I take here an adversarial approach.

Quitting is a reversible decision, e.g. storing the gun away somewhere, is cost-less and entails no risk of detection. In this sense, quitting is like *concealment*. By concealing the gun, the owner does not reap the benefits, x_t , but he also avoids the risk, $\rho\bar{R}$.

First, observe that the regulator can never induce self-reporting by the agent for $R_t > 0$, since concealment is always strictly preferred. I therefore focus on the case of $\underline{R} < 0$. Although I don’t derive the optimal policy in this case, I show that if $\underline{R} < 0$, i.e. the regulator has the power to operate a buy-back program, the basic force that drives the results of Section 2 remains.

Consider a simple case in which $x_l = \alpha_l = 0$. It is immediate to see that if $\underline{R} > 0$, no policy can induce any self-reporting; concealment forever is always possible and a strictly preferable option. Suppose then that $\underline{R} < 0$, i.e. the regulator has the power to run a buyback program.

Consider first a static program that offers $R_t = \underline{R}$ at all t . Since $\underline{R} \leq 0$, $x_l = \alpha_l = 0$, the agent immediately self-reports when he transitions to the low state. Now, what does an agent in the high state do? First, we can see that unless he immediately reports, he never conceals the weapon — the environment is stationary so if he conceals the weapon at some point in the high state, then he *always* conceals in the high state. But, he can improve on this by immediately self-reporting for \underline{R} . Thus, either the agent in the high state immediately reports or he never conceals. But, since he is immediately self-reporting in the low state, the analysis is exactly the same as without concealment! In particular, the agent finds it optimal to self-report in the high state if and only if $x_h \leq \rho\bar{R} - (\rho + r)\underline{R}$ and otherwise he operates without concealment.

On the other hand, consider an extreme, one-time buy-back in which, for some $T > 0$, the regulator sets $R_T = \underline{R}$ and for any $t \neq T$ the regulator sets $R_t = \bar{R}$. At any time after t , an agent in state x_l conceals the gun, but no longer has the option to receive \underline{R} (as in the static program). So if, at time T , the high type chooses not to take advantage of the

amnesty program, his payoff is $\frac{x_h - \rho \bar{R}}{\rho + r + \lambda}$. The agent prefers to self-report at T whenever

$$-\underline{R} \geq \frac{x_h - \rho \bar{R}}{\rho + r + \lambda} \implies \rho \bar{R} - \underline{R}(\rho + r + \lambda) \geq x_h$$

Combining this with our finding from the static program, we see that whenever

$$x_h - \rho \bar{R} \in (-(\rho + r)\underline{R}, -(\rho + r + \lambda)\underline{R}),$$

the once-and-for-all policy will induce agents in the high state to report at T , while the static policy never induces agents in the high state to report.

I conclude then that allowing for cost-less quitting leaves the basic force studied in the paper in tact, as long as $\underline{R} < 0$. If quitting or concealing is costly, then the force studied in the model continues to be relevant even if $\underline{R} \geq 0$ (depending on the size of the cost, for instance an infinite cost puts us back exactly in the model of Section 2).

A.3 Entry Deterrence

In this section, I allow agents to decide whether to begin committing crime, maintaining the assumption that the regulator cannot observe the time at which crime is committed.

Model. To formally allow for a decision to enter crime, I study exactly the model of Section 2 but introduce a third state for values, x_{hh} , which is higher than x_h . This state x_{hh} is so high that agents can be neither induced to self-report nor deterred from entering using any policy $(R_t)_{t \geq 0}$. Upon arrival to the model, agents make a once-and-for-all decision whether to begin committing crime. It is also important now to allow for arrival in both states x_h and x_{hh} (as in the baseline model of Section 2, allowing for arrival in state x_l does not change results). To this end, let μ be a time independent arrival distribution across states $\{x_{hh}, x_h\}$.⁴²

For simplicity, I study a case in which $x_l = 0$ and $\underline{R} = 0$.

Assumption 1. $\underline{R} = x_l = 0$.

Agents can transition only from x_{hh} to x_h or from x_h to x_l . As before, transitions from x_h to x_l occur at rate λ and transitions from x_{hh} to x_h occur at rate λ_{hh} . I assume the following on the new features of the model,

Assumption 2. $x_{hh} > \frac{(\rho + r + \lambda_{hh})}{\rho + r} \rho \bar{R}$,

⁴²Note that if μ put probability 1 on x_h , or equivalently $x_{hh} = x_h$, then the optimal policy would be static. If $R_t = \bar{R}$, then giving agents the option to not begin committing crime in the first place is like offering a once-and-for-all amnesty. If this does not deter x_h types from entering, then no self-reporting policy can induce self-reporting by x_h types after they have entered. So, either the static policy $R_t = \bar{R}$ for all t deters both x_h and x_l types from entering, or it does not and then the static policy $R_t = \underline{R}$ for all t is optimal — this policy at least induces immediate reporting by low types.

Assumption 3. $x_h \in (\rho\bar{R}, (\rho + r + \lambda)\frac{\rho}{\rho+r}\bar{R})$

If the first assumption fails, then setting $R_t = 0$ is optimal. This deters all agents from entering. If the second assumption fails, then setting $R_t = 0$ is optimal for the same reasons as in Theorem 2.

Analysis. It is immediate to see that if μ put probability 1 on x_{hh} , the analysis of the model proceeds unchanged by replacing arrival with arrival+transition to state x_h in the formulation of Section 2.

A more difficult but realistic setting allows μ to put positive probability on both x_{hh} and x_h . In designing her policy, the regulator must take account of a natural and important trade-off between (i) using of self-reporting to entice agents who entered in state x_{hh} but have transitioned to state x_h and (ii) shutting down self-reporting to ensure that state x_h agents do not report. As before, it is possible to induce reporting by x_l agents everywhere, so they do not pose any new difficulty in this environment. While I do not solve for the general optimal policy, I argue that a version of the cyclical policy in Theorem 2 strictly improves the regulator's value over static policies.

The assumptions made on x_{hh} and λ_{hh} guarantee that:

- $R_t = \bar{R}$ for all t deters agents in state x_h from entering but not agents in state x_{hh} (first and second requirement)
- agents in state x_h can be induced to self-report, but not with a static policy (second requirement)

As long as both $\mu_0(x_{hh}) > 0$ and $\mu_0(x_h) > 0$, a positive mass of x_{hh} agents enter and eventually transition to state x_h . In this case, the trade-off between entry deterrence and ex post detection comes to the fore.

I claim that there exists a pair $(T, v) \in \mathbb{R}_+ \times [\underline{R}, \bar{R}]$ such that the following hold,

$$\begin{aligned} & \frac{x_h}{\rho + r + \lambda_h}(1 - e^{-(\rho+r+\lambda_h)T}) - \frac{\rho\bar{R}}{\rho+r}(1 - e^{-(\rho+r)T}) - ve^{-(\rho+r)T} \leq -v \quad (\text{Detect}) \\ \max_{S \leq T} & \left\{ \frac{x_h}{\rho + r + \lambda_h}(1 - e^{-(\rho+r+\lambda_h)S}) - \frac{\rho\bar{R}}{\rho+r}(1 - e^{-(\rho+r)S}) - ve^{-(\rho+r)S} \right\} \leq 0 \quad (\text{Deter}) \end{aligned}$$

The first equation is a generalization of the one-shot incentive compatibility constraint Equation (I) that guarantees that an agent prefers to report when $R_t = v$ rather than wait to report at $t + T$ for $R_{t+T} = v$. The second guarantees that no agent who arrives in state x_h would decide to enter and become a criminal. Because $x_h < (\rho + r + \lambda)\bar{R}$, we know that there exist (T, v) which satisfy the first equation.

I argue now that there exists some (T, v) that satisfies both equations. The largest v for which there exists $T \in [0, \infty]$ that satisfies the first equation is $-v = \frac{x_h}{\rho+r+\lambda_h} - \frac{\rho\bar{R}}{\rho+r}$, which

satisfies the first equation at $T = 0$. Plugging into the left-hand side of the second equation, we get:

$$\begin{aligned}
& \max_{S \leq T} \left\{ \frac{x_h}{\rho + r + \lambda_h} (1 - e^{-(\rho+r+\lambda_h)S}) - \frac{\rho \bar{R}}{\rho + r} (1 - e^{-(\rho+r)S}) - v e^{-(\rho+r)S} \right\} \\
&= \max_{S \leq T} \left\{ \frac{x_h}{\rho + r + \lambda_h} (1 - e^{-(\rho+r+\lambda_h)S}) - \frac{\rho \bar{R}}{\rho + r} (1 - e^{-(\rho+r)S}) + \left(\frac{x_h}{\rho + r + \lambda_h} - \frac{\rho}{\rho + r} \bar{R} \right) e^{-(\rho+r)S} \right\} \\
&= \max_{S \leq T} \left\{ \frac{x_h}{\rho + r + \lambda_h} (1 - e^{-(\rho+r+\lambda_h)S} - e^{-(\rho+r)S}) - \frac{\rho \bar{R}}{\rho + r} \right\} \\
&\leq \frac{x_h}{\rho + r + \lambda_h} - \frac{\rho \bar{R}}{\rho + r} \leq 0
\end{aligned}$$

where the last inequality follows by our assumption on x_h . This shows that there exists a pair (T, v) satisfying both equations. Applying Lemma 1, the policy in Figure 5 guarantees

- any agent who enters in state x_h or x_{hh} immediately reports when in state x_l
- x_h type agents are always deterred from entering, by the Equation (Deter)
- x_h types report when $R_t = v$, by the second equation Equation (Detect)

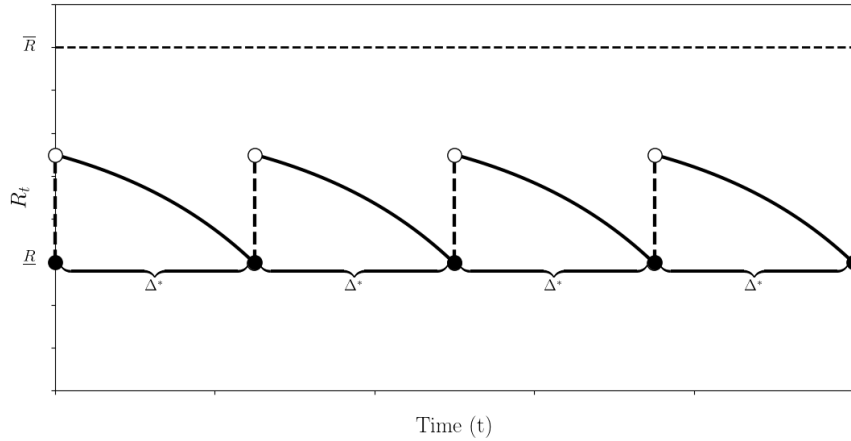


Figure 5: A cyclical policy satisfying Equation (Deter) and Equation (Detect).

Consider now a static policy, $R_t = c$ for all t . An agent in state x_i reports if and only if $x_i - (\rho \bar{R} - (\rho + r)c) \leq 0$. As long as $c < \frac{\rho}{\rho+r} \bar{R}$ an agent in state $x_l = 0$ will immediately report. For an agent in state x_h , his value to waiting until state x_l to report is $\frac{x_h - \rho \bar{R} - \lambda c}{\rho + r + \lambda_h}$. So, x_h types can be deterred from entering while at the same time inducing immediate reporting by the x_l types as long as $\frac{x_h - \rho \bar{R}}{\lambda} < c < \frac{\rho}{\rho+r} \bar{R}$. However, since $x_{hh} > x_h > \rho \bar{R}$, no agent in state x_h who has entered in state x_{hh} ever self-reports. Therefore, the cyclical policy represents a strict improvement over any static policy:

- it deters agents in state x_h and x_l from entering, (like a static policy)
- induces immediate reporting in state x_l , (like a static policy)
- induces occasional reporting by agents in state x_h (*unlike* a static policy)

Randomization. While this policy does improve the regulator’s value relative to static policies, it is not in general optimal. At an intuitive level, randomization appears critical to achieve the optimal value because by randomizing, the regulator relaxes her (Deter) constraint. For instance, in the cyclical policy, agents arriving in state x_h at the very beginning of a cycle have a much lower value to entering than do agents x_h arriving in the middle or end of a cycle. Randomizing allows the regulator to spread this deterrence more uniformly over the cycle.

It is straightforward to find numerical examples that can achieve the same level of deterrence as in the cyclical policy above — agents in states x_h and x_l never enter — but improve the regulator’s value. For instance, consider the Poisson random policies studied in Appendix A.1. Any such policy that satisfies the agent’s incentive compatibility condition *also* deters agents in state x_h from entering, because the arrival of an amnesty is i.i.d. over time.

The description of an optimal policy in this setting with fully general random policies is very interesting, but left open.

A.4 Arrival Time

To further investigate the role of arrival time in policy design, I take up the analysis of the model when the regulator can use some information about arrival time. There are two relevant versions, both in which the regulator faces a single agent who (i) arrives at time 0 in a very high state x_{hh} and transitions to x_h at Poisson rate γ (without ever transitioning back) or (ii) arrives in state x_h at exponential rate γ .

As long as x_{hh} is sufficiently large, the mechanics of these models are the same. Therefore, I only provide the formal results for the first version when x_{hh} is sufficiently large. Suppose that the regulator observes the arrival time of the agent (normalized to 0), but that the agent arrives in state x_{hh} . The agent transitions at rate γ from state x_{hh} to state x_h and never transitions back to state x_{hh} . The state x_{hh} is to be interpreted as a state that self-reporting is never optimal for the agent, so we need an assumption on the size of x_{hh} to make this so:

Assumption 4. $\frac{x_{hh}}{\rho+r+\gamma+\lambda} > \frac{\rho\bar{R}}{\rho+r} - \underline{R}$

This is a sufficient condition that ensures that no self-reporting policy can ever induce the agent to report when in state x_{hh} . Let $\mathbf{1}_{j,T}$ be an indicator that an agent has arrived before T , has not been caught or reported by T , and is in state j at T . Given a reporting policy,

R , let $\mathbb{M}(R)$ be the joint distribution of $(\mathbf{1}_{x_j,T}, \mathbf{1}_{x_l,T})_{T \geq 0}$. Then, we have:

$$V_\gamma(R) := -\mathbb{E}_{\mathbb{M}(R)} \left[\int_0^\infty e^{-rs} (\alpha_{hh} \mathbf{1}_{x_{hh},s} + \alpha_h \mathbf{1}_{x_h,s} + \alpha_l \mathbf{1}_{x_l,s}) ds \right]$$

and as before, the regulator chooses:

$$V_\gamma^* = \sup_R V_\gamma(R)$$

It turns out that the steps for proving Proposition E.1 apply with little adjustment to V_γ^* , as long as $\gamma < \rho$. In particular, I have the following proposition:

Proposition A.1. *Suppose that Assumption 4 holds and $\gamma < \rho$. Then the optimal policy in Theorem 2 remains limit-optimal.*

When $\gamma > \rho$, the regulator believes that the agent arrives relatively quickly after time 0. In an extreme case when γ gets very large, the regulator knows with near certainty that the agent is in state x_h near 0. In this case, the regulator is willing to trade a long future period without reporting by the high type agent for more frequent reporting by the high type near time 0.

A.5 Generalized Multi-Agent Setting

In this section, I generalize the results of Section 6 to a case in which the flow value x_t can take more than two values. To enrich the scope of the analysis, x_t is a continuous-time Markov chain with J states $x^1 < x^2 < \dots < x^J$. From state x^j , the process x_t can transition only to x^{j-1} and does so at rate λ^j . As convention, let $\lambda^1 := 0$. The process in Section 6 is a special case when there are just two states, $x^1 < x^2$. I maintain the assumptions made in Section 6.1 but generalize the first to avoid non-generic issues,

Assumption: For all j , $x^j - (\rho + \lambda^j)\bar{R} \neq 0$ and $x_j - (\rho + \frac{\lambda^j}{2})\bar{R} \neq 0$. For any criminal organization, let $\tau^j := \inf\{t | x_t \leq x_j\}$. The proofs of the results in this section are intuitive but tedious in places. As such, they are relegated to Online Appendix F.

A.5.1 Results: Unconstrained p

It is useful to define now two values,

$$\begin{aligned} \underline{J}^* &= \max\{j | \forall k \leq j, x^k - (\rho + \frac{\lambda^k}{2})\bar{R} < 0\} \\ \bar{J}^* &= \max\{j | \forall k \leq j, x^k - (\rho + \lambda^k)\bar{R} < 0\} \end{aligned}$$

and if \underline{J}^* and/or \overline{J}^* does not exist then set it equal to 0. As we will see, \underline{J}^* and \overline{J}^* represent the minimum amount of unraveling in some equilibrium under various restrictions on the policies at the regulator's disposal.

As in Section 6, define (R^*, p^*) as,

- $R_t^* = 0$ for all t
- $p_{1,t}^* = p_{2,t}^* = 0$

The following proposition generalizes Proposition 2.

Proposition A.2. *Fix any (R, p) . Then there exists a pure strategy equilibrium (τ_1^*, τ_2^*) such that $\tau_i^* \geq \tau^{\overline{J}^*}$. Under the policy (R^*, p^*) , any equilibrium $m^* = (m_1^*, m_2^*)$ is such that*

$$\mathbb{P}_{m^*}(\min\{\tau_1^*, \tau_2^*\} \leq \tau^{\overline{J}^*}) = 1.$$

The proof is given in Appendix F. The first part of the proposition states that regardless of the p and R the regulator uses, there is always an equilibrium in which agents do not report until the state has reached $x^{\overline{J}^*}$ or lower. The second part states that the policy (R^*, p^*) guarantees that in *every* equilibrium, agents immediately report by the time they reach state $x^{\overline{J}^*}$.

A.5.2 Results: Simple Policies

In this section, I generalize the results of Appendix A.5.2 to the more general x_t process described above. In particular, as in Appendix A.5.2, I restrict the regulator to simple policies i.e. $p_{i,t} = \frac{1}{2}$ for each i and t . Static simple policies cannot, in general, achieve the performance of the policy in Proposition A.2 — unraveling can be guaranteed at most to \underline{J}^* , not \overline{J}^* .

Proposition A.3. *Fix any simple static policy, $R_t = v$ for all t . Then there exists a pure strategy equilibrium (τ_1, τ_2) s.t. $\tau_i \geq \tau^{\underline{J}^*}$ for each i .*

This proposition states that there exist pure strategy equilibria in which unraveling stops at state \underline{J}^* . The transition from state $\underline{J}^* + 1$ to state \underline{J}^* is sufficiently slow that i prefers to wait until state \underline{J}^* to report, given that $-i$ plans to do the same.

I proceed to generalize Proposition 4. Recall that the policy $R^*(\epsilon, T)$ is defined as:

1. Let $R_t^* := \overline{R} - \epsilon$ for a length of time T
2. Let $R_t^* := \overline{R} - \epsilon - t$ for a length of time $\overline{R} - \epsilon$
3. Return to (1) and repeat

The following proposition generalizes Proposition 4,

Proposition A.4. *There exists $R^*(\epsilon, T)$ such that if $m^* = (m_1^*, m_2^*)$ is an equilibrium, then*

$$\mathbb{P}_{m^*}(\min\{\tau_1, \tau_2\} \leq \tau^{\bar{J}^*}) = 1.$$

The proof of the proposition is given in Appendix F, which proceeds by iteratively deleting strategies. The idea is the same as in Proposition 4, except in this case the unraveling does not stop until it reaches state $x^{\bar{J}^*}$, rather than x_h .

A corollary of these results is that, under agent-preferred equilibrium selection, $R^*(\epsilon, T)$ achieves the same outcome as the non-simple policy (R^*, p^*) . Further, if $\underline{J}^* < \bar{J}^*$, the set of static simple policies *cannot* achieve this performance.

B. Proof of Proposition 1

Proof of Proposition 1: Because R^v is constant and hence continuous, Theorem 3 in Shiryaev (2007) can be applied to show that there exists some $D \subset \{x_h, x_l\}$ such that⁴³

1. $\tau_v^* := \inf\{t \geq t_0 | x_t \in D\}$,
2. If τ is any other optimal stopping time for the agent, then $\mathbb{P}(\tau_v^* \leq \tau) = 1$.

It is therefore without loss of generality for the regulator to restrict to recommendation policies A such that $A_t(x) = A_s(x)$ for all $t, s \geq 0$ and $x \in \{x_h, x_l\}$, since these induce all stopping times of the form τ_v^* .

To prove the lemma, it is thus sufficient to argue that $\tau_{\underline{R}}^* \leq \tau_v^*$ for all $v \geq \underline{R}$. Given the characterization of τ_v^* described above, the only possibilities are

1. $\tau_v^* = \tau^{*,0} := 0$
2. $\tau_v^* = \tau^{*,\infty} := \infty$
3. $\tau_v^* = \tau^{*,l} := \inf\{t | x_t = x_l\}$

The agent's value for $\tau^{*,\infty}$ is independent of v . The agent's values for $\tau^{*,0}$ and $\tau^{*,l}$ are

$$\begin{aligned} \mathbb{E}[W(x, t, \tau^{*,0}; R^v)] &= -v \\ \mathbb{E}[W(x, t, \tau^{*,l}; R^v)] &= \mathbf{1}_{x=x_h} \left(\frac{x_h}{\rho + r + \lambda} + \frac{x_l - \rho \bar{R}}{\rho + r} - \frac{v\lambda}{\rho + r + \lambda} \right) + \mathbf{1}_{x=x_l}(-v) \end{aligned}$$

To conclude that $\tau_{\underline{R}}^* \leq \tau_v^*$, observe that decreasing v increases $\mathbb{E}[W(x, t, \tau^{*,0}; R^v)]$ by weakly more than $\mathbb{E}[W(x, t, \tau^{*,l}; R^v)]$. Similarly, decreasing v weakly increases $\mathbb{E}[W(x, t, \tau^{*,l}; R^v)]$ but has no effect on $\mathbb{E}[W(x, t, \tau^{*,\infty}; R^v)]$. Decreasing v can therefore only induce a switch from $\tau^{*,\infty}$ to one of the other two, or from $\tau^{*,l}$ to $\tau^{*,0}$. The conclusion follows. \square

⁴³Application of the theorem in Shiryaev (2007) requires a minor re-casting of the stopping problem presented here. In particular, the state space must be expanded to account for the accumulating value. This formulation is straightforward and therefore omitted.

C. Proof of Lemma 1

Proof of Lemma 1: First, given (R, A) , consider an alternative policy (\hat{R}, \hat{A}) s.t.

- $\hat{R}_t = R_t$ if $A_t(x_h) = 1$
- $\hat{R} = \bar{R}$ if $A_t(x_h) = 0$
- $\hat{A}_t(x_h) = \hat{A}_t(x_l) = 1$ if $A_t(x_h) = 1$
- $\hat{A}_t(x_h) = A_t(x_l) = 0$ if $A_t(x_h) = 0$

Observe that \hat{A} is obedient and since A, R are measurable, then so are \hat{A} and \hat{R} , so that $(R, A) \in \mathcal{M}$. I now generate a new policy, (\tilde{R}, \tilde{A}) , such that $\hat{A}_t(x_h) = \tilde{A}_t(x_h)$ and $\tilde{A}_t(x_l) = 1$ for all t . The definition of W^* and the assumption that $0 \leq x_l < \rho\bar{R} - (\rho + r)\underline{R}$ imply that $W^*(x_l, t; \hat{R}) \in [-\bar{R}, -\underline{R}]$ for all t . The penalty process \tilde{R} is defined by

- $\tilde{R}_t = -W^*(x_l, t; \hat{R})$ if $\hat{A}(x_h, t) = 0$
- $\tilde{R}_t = R_t$ if $\hat{A}(x_h, t) = 1$

I argue now that \tilde{R} is measurable. Let $\underline{T}^h(t) := \inf \left\{ s \in \{s | \tilde{A}_s(x_t) = 1\} \cap [t, \infty) \right\}$ — since \tilde{A} is measurable so is $\underline{T}^h(t)$. Then, observe that

$$W^*(x_l, t; \hat{R}) = \frac{x_l - \rho\bar{R}}{\rho + r} (1 - e^{-(\rho+r)(\underline{T}^h(t)-t)}) - e^{-(\rho+r)(\underline{T}^h(t)-t)} R_{\underline{T}^h(t)}$$

which is measurable since $\underline{T}^h(t)$ is.

Consider any stopping policy τ under this this new policy. Then, we have:

$$W(x, t, \tau; \tilde{R}) = \mathbb{E} \left[\int_0^{\tau \wedge \tau_\rho} e^{-r(s-t)x_s} ds - e^{-r(\tau \wedge \tau_\rho - t)} \left(\mathbf{1}_{\tau_\rho \leq \tau} \bar{R} + \mathbf{1}_{\tau < \tau_\rho} \left(\mathbf{1}_{\tilde{A}(x_h, t) = 1} R_\tau + \mathbf{1}_{\tilde{A}(x_h, t) = 0} W^*(x_l, \tau; \tilde{R}) \right) \right) \right]$$

Now, let $\sigma := \tau \mathbf{1}_{A_\tau(x_h) = 0} + \infty (1 - \mathbf{1}_{A_\tau(x_h) = 0})$. Then, this expression can be written:

$$\begin{aligned} W(x, t, \tau; \tilde{R}) &= \mathbb{E} \left[\int_0^{\tau \wedge \tau_\rho \wedge \sigma} e^{-r(s-t)x_s} ds - e^{-r(\tau \wedge \tau_\rho \wedge \sigma - t)} \left(\mathbf{1}_{\tau_\rho \leq \tau \wedge \sigma} \bar{R} + \mathbf{1}_{\tau < \tau_\rho \wedge \sigma} R_\tau + \mathbf{1}_{\sigma \leq \tau_\rho \wedge \tau} W^*(x_l, \sigma; \tilde{R}) \right) \right] \\ &\leq \mathbb{E} \left[\int_0^{\tau \wedge \tau_\rho \wedge \sigma} e^{-r(s-t)x_s} ds - e^{-r(\tau \wedge \tau_\rho \wedge \sigma - t)} \left(\mathbf{1}_{\tau_\rho \leq \tau \wedge \sigma} \bar{R} + \mathbf{1}_{\tau < \tau_\rho \wedge \sigma} R_\tau + \mathbf{1}_{\sigma \leq \tau_\rho \wedge \tau} W^*(x_\sigma, \sigma; \tilde{R}) \right) \right] \end{aligned}$$

where the second line is simply a result of the fact that $x_l \leq x_\sigma$. But then this implies that

$$W(x, t, \tau; \tilde{R}) \leq W^*(x, t; \hat{R})$$

Conversely, by using the strategy $\tau := \{t | A_t(x_h) = 1\}$, an agent guarantees himself $W^*(x_i, t; \hat{R})$, and so we conclude that $W^*(x_i, t; \hat{R}) = W^*(x_i, t; \tilde{R})$. Then by the definition of (\tilde{R}, \tilde{A}) , $\tau := \inf\{t | \tilde{A}_t(x_t) = 1\}$ is optimal so \tilde{A} is an obedient recommendation strategy. This completes the proof. \square

D. Proof of Theorem 1

I first prove an intermediate result, for the case in which $x_l = \alpha_l = 0$. Since there is no ambiguity in this case, let $x := x_h$.

Proposition D.1. *Suppose $x_l = \alpha_l = 0$. Then,*

$$\Theta^* = \left\{ (\rho, r, \lambda, \bar{R}, \underline{R}, x) \mid (\rho + r + \lambda) \frac{\rho \bar{R} - (\rho + r) \underline{R}}{\rho + r} \geq x > \rho \bar{R} - (\rho + r) \underline{R} \right\}$$

Proof. By Proposition 1, it is without loss of generality to suppose that, for static penalty policies R^v , the regulator's recommendation is constant i.e. $A_t(x) = A_s(x)$ for each $t, s \geq 0$ and $x \in \{x_h, x_l\}$. An agent's stopping times is then only one of the following three,

1. $\tau^{*,0} := 0$
2. $\tau^{*,\infty} := \infty$
3. $\tau^{*,l} := \inf\{t | x_t = x_l\}$

Observe that, for the case of x_t considered here, the regulator is indifferent between policies $\tau^{*,\infty}$ and $\tau^{*,l}$.⁴⁴ By Proposition 1 once more, I need only show a strict improvement over the static penalty policy $R^{\underline{R}}$ if I want to demonstrate that static policies can be strictly improved.

I first show that $\Theta^* \subseteq \{(\rho, r, \lambda, \bar{R}, \underline{R}, x) \mid (\rho + r + \lambda) \frac{\rho \bar{R} - (\rho + r) \underline{R}}{\rho + r} \geq x_h > \rho \bar{R} - (\rho + r) \underline{R}\}$. Fix an arbitrary set of parameters, $\theta \in \Theta^*$. To show this, I compute the value to the agent under $R^{\underline{R}}$ for any of his three possible optimal stopping times of the agent. The values under penalty policy $R^{\underline{R}}$ for an agent arriving at t_0 in state x are

$$\begin{aligned} W(x, \tau^{*,l}, t_0; R^{\underline{R}}) &= \mathbb{E}_{x_0} \left[\int_0^{\tau^{*,l}} x_t e^{-(\rho+r)t} - \frac{\rho}{\rho+r} \bar{R} (1 - e^{-(\rho+r)\tau^{*,l}}) - e^{-(\rho+r)\tau^{*,l}} \underline{R} \right] \\ &= \mathbf{1}_{x=x_h} \left(\frac{x_h - \rho \bar{R} - \lambda \underline{R}}{\rho + r + \lambda} \right) + \mathbf{1}_{x=x_l} (-\underline{R}) \\ W(x_h, \tau^{*,0}, t_0; R^{\underline{R}}) &= -\underline{R} \\ W(x_h, \tau^{*,\infty}, t_0; R^{\underline{R}}) &= \mathbf{1}_{x=x_h} \left(\frac{x_h}{\rho + r + \lambda} \right) - \frac{\rho \bar{R}}{\rho + r} \end{aligned}$$

⁴⁴In particular, the low flow state generates no direct loss, and there is no chance of transitioning out of this state.

Suppose first that $x_h - (\rho\bar{R} - (\rho + r)\underline{R}) \leq 0$. Then,

$$\max\{W(x_h, \tau^{*,0}, t_0), W(x_h, \tau^{*,l}, t_0)\} \leq W(x_h, \tau^{*,0}, t_0)$$

As a consequence, the recommendation policy $(A_t)_{t \geq 0}$ with $A_t(x) = 1$ for all t, x is an obedient recommendation policy. The regulator achieves first best with $R^{\underline{R}}$ and $A_t(x) = 1$ for all t, x . This implies that $\Theta^* \subseteq \{(\rho, r, \lambda, \bar{R}, \underline{R}, x) | x_h - (\rho\bar{R} - (\rho + r)\underline{R}) > 0\}$.

Suppose now that $x_h > (\rho + r + \lambda)(\frac{1}{\rho+r})(\rho\bar{R} - (\rho + r)\underline{R})$. The agent's value for $\tau^{*,\infty}$ is

$$\begin{aligned} W(x, \tau^{*,\infty}, t_0; R^{\underline{R}}) &= \mathbb{E} \left[\int_0^\infty x_t e^{-(\rho+r)t} dt - \frac{\rho}{\rho+r} \bar{R} \right] \\ &= \mathbf{1}_{x=x_h} \frac{x_h}{\rho+r+\lambda} - \frac{\rho\bar{R}}{\rho+r} \end{aligned}$$

When $x = x_h$, the assumption that $x_h > (\rho + r + \lambda)(\frac{1}{\rho+r})(\rho\bar{R} - (\rho + r)\underline{R})$ implies that this expression is strictly larger than $-\underline{R}$. Because of this, no recommendation policy with $A_t(x_h) = 1$ for some t can be obedient. This implies that $\Theta^* \subseteq \{(\rho, r, \lambda, \bar{R}, \underline{R}, x) | (\rho + r + \lambda)(\frac{\rho\bar{R} - (\rho+r)\underline{R}}{\rho+r}) \geq x_h\}$. Combining this with our finding above that $\Theta^* \subseteq \{(\rho, r, \lambda, \bar{R}, \underline{R}, x) | x_h - (\rho\bar{R} - (\rho + r)\underline{R}) > 0\}$, the first inclusion is shown.

I now show $\left\{ (\rho, r, \lambda, \bar{R}, \underline{R}, x) | (\rho + r + \lambda)(\frac{\rho\bar{R} - (\rho+r)\underline{R}}{\rho+r}) \bar{R} \geq x > \rho\bar{R} - (\rho + r)\underline{R} \right\} \subseteq \Theta^*$. Let θ be an arbitrary element on the left-hand side. Consider the policy $R^{\underline{R}}$. Using the equations above and the assumption $x_h > \rho\bar{R} - (\rho + r)\underline{R}$,

$$W(x_h, \tau^{*,l}, t_0) - W(x_h, \tau^{*,0}, t_0) = \frac{x - \rho\bar{R} - \lambda\underline{R}}{\rho + r + \lambda} + \underline{R} > 0$$

In this case then, the regulator receives his *worst possible* payoff; no agent ever reports until reaching the low state $x_l = 0$. Thus, all I need to do to conclude the proof is demonstrate a policy which induces a positive mass of high types to report.

To this end, consider a *one-time policy*: $\tilde{R}_t = \mathbf{1}_{t=T}\underline{R} + (1 - \mathbf{1}_{t=T})\bar{R}$ for some $T > 0$ and $A_T(x_l) = A_T(x_h) = 1$. Then, observe that,

$$W(x_h, \tau^{*,\infty}, T) = \frac{x_h}{\rho + r + \lambda} - \frac{\rho\bar{R}}{\rho + r} \leq -\underline{R} = W(x_h, \tau^{*,0}, T)$$

where the inequality follows by assumption that $(\rho + r + \lambda)(\frac{\rho\bar{R} - (\rho+r)\underline{R}}{\rho+r}) \geq x_h$. Thus, the recommendation $A_t(x) = 1$ if and only if $t = T$ is obedient. Since $T > 0$, this policy induces a *strictly positive* mass of high types to stop by T , generating a strict improvement of the regulator's value over any static policy. So $\{(\rho, r, \lambda, \bar{R}, \underline{R}, x) | (\rho + r + \lambda)(\frac{\rho\bar{R} - (\rho+r)\underline{R}}{\rho+r}) \geq x > \rho\bar{R} - (\rho + r)\underline{R}\} \subseteq \Theta^*$ and the proof is concluded by combining this with the reverse inclusion. \square

Proof of Theorem 1: Suppose now that $x_l > 0$ or $\alpha_l > 0$. Suppose first that $x_l > \rho\bar{R} - (\rho + r)\underline{R}$. Then, $W(x_l, \tau^{*,\infty}, t; R) > 0$ for any t, R where $\tau^{*,\infty} = \infty$. In that case, the only obedient recommendation is $A_t(x_h) = A_t(x_l) = 0$ for all t , in which case the policy R has no effect on behavior. Because of this, $\Theta^* \cap \{(\rho, r, \lambda, \bar{R}, \underline{R}, x) | x_l > \rho\bar{R} - (\rho + r)\underline{R}\} = \emptyset$.

Suppose now that $x_l \leq \rho\bar{R} - (\rho + r)\underline{R}$. Apply Lemma 1 to see that $V^* = \sup_{(R,A) \in \mathcal{M}} V(R, A) = \sup_{(R,A) \in \mathcal{L}} V((R, A))$. Since the right hand side is independent of α_l , it is without loss of generality to prove the result for the $x_l \geq 0$ but $\alpha_l = 0$. To this end, set $\tilde{x}_h = x_h - x_l$, $\tilde{x}_l = 0$ and $\tilde{R} = \bar{R} - \frac{x_l}{\rho}$. An agent's value for a stopping time can then be re-written,

$$\begin{aligned} W(x, t_0, \tau; R) &= \mathbb{E} \left[\int_0^\tau e^{-(\rho+r)t} x_t dt - (1 - e^{-(\rho+r)\tau}) \frac{\rho}{\rho+r} \bar{R} - e^{-(\rho+r)\tau} R_{\tau+t_0} \right] \\ &= \mathbb{E} \left[\int_0^{\tau \wedge \tau_l} e^{-(\rho+r)t} (x_h - x_l) dt + \int_0^\tau e^{-(\rho+r)t} (x_l) dt - (1 - e^{-(\rho+r)\tau}) \frac{\rho}{\rho+r} \bar{R} - e^{-(\rho+r)\tau} R_{\tau+t_0} \right] \\ &= \mathbb{E} \left[\int_0^\tau e^{-(\rho+r)t} \tilde{x}_t dt - (1 - e^{-(\rho+r)\tau}) \frac{\rho \tilde{R}}{\rho+r} - e^{-(\rho+r)\tau} R_{\tau+t_0} \right] \end{aligned}$$

So, an agent's value for a stopping time is the same across parameterizations $(\rho, r, \lambda, \bar{R}, \underline{R}, \tilde{x}_h, \tilde{x}_l)$ and $(\rho, r, \lambda, \bar{R}, \underline{R}, x_h, x_l)$. The result then follows from the application of Proposition D.1 to the parameterization $\theta = (\rho, r, \lambda, \bar{R}, \underline{R}, \tilde{x}_h, \tilde{x}_l)$. \square

E. Proof of Theorem 2

I first prove an intermediate proposition, which is the theorem restricted to the case of $x_l = \alpha_l = 0$. Some of the less important lemmas whose proofs are not particularly instructive are relegated to Online Appendix G.

Proposition E.1. *Suppose $\alpha_l = x_l = 0$. If $\theta \in \Theta^*$, an optimal policy, (R^*, A^*) , is:*

- $R_{nT^*+T_0}^* = \underline{R}$ for $n \in \mathbb{N}$ for some $T_0 \geq 0$
- $R_t^* = \bar{R}$ otherwise
- $A_t^*(x_h) = 1$ if and only if $t \in \{T_0, nT^*\}_{n \in \mathbb{N}}$
- $A_t^*(x_l) = 1$ for all $t \geq 0$

I build to this proposition through a number of lemmas. Below is a roadmap with a description of each result:

- **Lemma E.1:** The first lemma shows how to compute the expectation of N_t^h , the number of agents in the high state at some time t .
- **Lemma E.2:** This lemma states that the regulator's problem can be approximated by a problem in which the regulator is restricted to policies such that for all t, s with $|t - s| < \epsilon$, we do not have $A_t(x_h) = A_s(x_h) = 1$. That is, the approximation restricts to the set of policies that does not induce reporting by the high type more than once in any ϵ interval. As ϵ shrinks to 0, the approximation error disappears.
- **Lemma E.3:** This lemma states that, if $\alpha_l = x_l = 0$, the only relevant information for the regulator's value contained in a policy R is the set of times at which the high type reports under his optimal stopping time. Thus, the regulator's problem can be solved by optimally choosing these times, subject to incentive compatibility conditions.
- **Lemma E.4:** This lemma states that a policy is incentive compatible if and only if it satisfies a one-shot deviation condition.
- **Lemma E.5:** This lemma argues that, except for an initial choice, the regulator's problem can be stated recursively where regulator chooses the next time and penalty level at which the high types will report and the state variable is a promised penalty which must be delivered immediately while respecting one-shot incentive compatibility.
- **Lemmas E.6 to E.9:** In these lemmas, I guess and verify an optimal policy to the dynamic program presented in Lemma E.5. Lemmas E.6 to E.8 prove some simple properties of the guess, while Lemma E.9 verifies the guess as optimal.

Among these, these most challenging is Lemma E.9, where the guess of the optimal policy is verified.

If a result requires $x_l = 0$, I will use x and x_h interchangeably. Let

$$T^h(t) := \sup\{s | 0 \leq s \leq t \text{ and } A_s(x_h) = 1\}$$

$T^h(t)$ is the elapsed time, at t , since high types last reported. Measurability of A implies measurability of $T^h(t)$. A few pieces of notation prove useful. Let:

$$F(T, T') := \alpha_h \int_T^{T'} \frac{1 - e^{-(\rho+\lambda)(t-T)}}{\rho + \lambda} e^{-r(t-T)} dt$$

$$C(T) := \frac{1 - e^{-(\rho+r+\lambda)T}}{\rho + r + \lambda} - \frac{\rho \bar{R}}{\rho + r} (1 - e^{-(\rho+r)T})$$

The first definition, F , is the regulator's loss from a period of time, T to T' , such that $A_T(x_h) = 1$ but $A_t(x_h) = 0$ for all $t \in (T, T')$. The second definition, $C(T)$, is part of the left-hand side of Equation (I).

Lemma E.1. *The discounted expected number of high types agents is:*

$$\mathbb{E} [e^{-rt} N_t^h] = \frac{1 - e^{-(\rho+\lambda)(t-T^h(t))}}{\rho + \lambda} e^{-rt}$$

Let \mathcal{M}^ϵ be the subset of policies $(R, A) \in \mathcal{M}$ such that for every t, s with $|t - s| < \epsilon$, we do not have $A_t(x_h) = A_s(x_h) = 1$. The regulator's ϵ problem is defined as:

$$V^{\epsilon,*} = \sup_{(R,A) \in \mathcal{M}^\epsilon} V(R) \quad (\mathcal{P}^\epsilon)$$

Lemma E.2. *The limit of $V^{\epsilon,*}$ as $\epsilon \rightarrow 0$ is V^* .*

I proceed to find a solution for $V^{\epsilon,*}$. The solution I find is independent of ϵ for ϵ sufficiently small and so it also solves V^* .

In what follows, I will identify a policy $(R, A) \in \mathcal{M}^\epsilon$ with the times at which $A_t(x_h) = 1$ as well as the penalties at such points. To this end, let $h = (T_i, R_i)_{i \leq N}$ be an arbitrary sequence with $N \in \mathbb{N} \cup \{\infty\}$ such that, $T_{i+1} - T_i \geq \epsilon$ and $R_i \in [\underline{R}, \bar{R}]$. Then (R^h, A^h) is the policy induced by $h = (T_i, R_i)_{i \leq N}$ with:

- $R_{T_i}^h = R_i$
- $R_t^h = \bar{R}$ otherwise
- $A_t^h(x) = 1$ if and only if $t = T_i$ for some i

Now, let \mathcal{H}^ϵ be the set of all sequences $h = (T_i, R_i)_{i \leq N}$ such that $(R^h, A^h) \in \mathcal{M}^\epsilon$. That is, the sequences in \mathcal{H}^ϵ induce penalties and associated recommendations which are incentive compatible. As convention, let $T_{-1} = 0$.

Lemma E.3. *Suppose $\alpha_l = x_l = 0$. Then*

$$V^{\epsilon,*} = \sup_{\mathcal{H}^\epsilon} - \sum_{i=0}^N e^{rT_{i-1}} F(T_{i-1}, T_i) - e^{-rT_N} F(T_N, \infty)$$

The next lemma is a manifestation of the one-shot deviation principle. The proof is standard and thus omitted.

Lemma E.4. *Suppose $x_l = 0$. Given any $h = (T_i, R_i)_{i \leq N}$ such that $T_{i+1} - T_i \geq \epsilon$,*

$$(R^h, A^h) \in \mathcal{M}^\epsilon \quad \iff$$

$$x_h \frac{1 - e^{-(\rho+r+\lambda)(T_{i+1}-T_i)}}{\rho+r+\lambda} - \frac{\rho\bar{R}}{\rho+r} (1 - e^{-(\rho+r)(T_{i+1}-T_i)}) - e^{-(\rho+r)(T_{i+1}-T_i)} R_{i+1} \leq -R_i$$

where $T_{i+1} = \infty$ if $N < \infty$.

The next lemma provides a way to compute the regulator's problem from a recursive equation.

Lemma E.5. *Suppose that $\frac{x_h}{\rho+r+\lambda} - \frac{\rho\bar{R} - (\rho+r)\underline{R}}{\rho+r} \leq 0$. Then suppose that $V^{D,\epsilon}$ is bounded and solves the following recursive equation on $[\underline{R}, \frac{\rho\bar{R}}{(\rho+r)} - \frac{x_h}{\rho+r+\lambda}]$:*

$$\begin{aligned} V^{D,\epsilon}(\ell) = & \sup_{T, \ell' \in [\underline{R}, \frac{\rho\bar{R}}{(\rho+r)} - \frac{x_h}{\rho+r+\lambda}]} -F(0, T) + e^{-rT} V^{D,\epsilon}(\ell') & (D) \\ & \text{s.t. } C(T) - e^{-(\rho+r)T} \ell' \leq -\ell \\ & \infty \geq T \geq \epsilon \end{aligned}$$

and suppose that $\tilde{\ell}(\ell)$ and $\tilde{T}(\ell)$ are optimal choices of ℓ' and T , respectively, which achieve value $V^{D,\epsilon}(\ell)$. Then

$$V^{*,\epsilon} = \sup_{T_0 \geq 0, R_0} -F(0, T_0) + e^{-rT_0} V^{D,\epsilon}(R_0) \quad (DI)$$

Recall that θ represents an arbitrary collection of parameters. Now, let $T^*(\ell)$ be the unique strictly positive solution to the equation (in T):

$$C(T) - e^{-(\rho+r)T} \underline{R} = -\ell \quad (I_\ell)$$

for any ℓ on $[\underline{R}, \frac{\rho\bar{R}}{\rho+r} - \frac{x_h}{\rho+r+\lambda}]$ if it exists and 0 otherwise. The next three lemmas establish some simple properties of T^* and F .

Lemma E.6. *Suppose $\theta \in \Theta^*$. Then, $T^*(\ell)$ is strictly increasing on $\ell \in [\underline{R}, \frac{\rho\bar{R}}{\rho+r} - \frac{x_h}{\rho+r+\lambda}]$*

Lemma E.7. *For any T, T' :*

$$\begin{aligned} & -F(0, T) - e^{-rT} F(0, T') - e^{-r(T+T')} \frac{F(0, T^*(\underline{R}))}{1 - e^{-rT^*}} \\ = & C_1 + C_2 \left[1 - e^{-(\rho+r+\lambda)T} + e^{-rT} (1 - e^{-(\rho+r+\lambda)T'}) + e^{-r(T+T')} \frac{1 - e^{-(\rho+r+\lambda)T^*(\underline{R})}}{1 - e^{-rT^*(\underline{R})}} \right] \end{aligned}$$

where $C_1 = -\frac{\alpha_h}{r(\rho+\lambda)}$ and $C_2 = \frac{\alpha_h}{(\rho+\lambda)(\rho+r+\lambda)}$.

Now, let $\tilde{V}^{D,\epsilon}$ be the value induced by policies $\tilde{T}(\ell) = T^*(\ell)$ and $\tilde{\ell}(\ell) = \underline{R}$. That is

$$\tilde{V}^{D,\epsilon}(\ell) := -F(0, T^*(\ell)) - e^{-rT^*(\ell)} \frac{F(0, T^*(\underline{R}))}{1 - e^{-rT^*(\underline{R})}}$$

Lemma E.8. Suppose that $\theta \in \Theta^*$. Then

$$-F(0, T) - e^{-rT} \frac{F(0, T^*(\underline{R}))}{1 - e^{-rT^*(\underline{R})}}$$

is decreasing in T for any $T \geq T^*(\underline{R})$. Furthermore, $\tilde{V}^{D, \epsilon}(\ell)$ is decreasing in ℓ .

Lemma E.9. Suppose that $\theta \in \Theta^*$ and $\epsilon < T^*(\underline{R})$. Then, the policies $\tilde{T}(\ell) = T^*(\ell)$ and $\tilde{\ell}(\ell) = -\underline{R}$ are optimal policies in Equation (D).

Proof. To show that $\tilde{T}, \tilde{\ell}$ is an optimal pair of functions with associated value $\tilde{V}^{D, \epsilon}(\ell)$, we must show that for any ℓ :

$$\begin{aligned} \tilde{V}^{D, \epsilon}(\ell) &= \sup_{T, \ell'} -F(0, T) + e^{-rT} \tilde{V}^{D, \epsilon}(\ell') \\ \text{s.t. } & C(T) - e^{-(\rho+r)T} \ell' \leq -\ell \end{aligned} \quad (*)$$

First, I argue that any choice of $T > T^*(\ell)$ can be strictly improved by setting $T = T^*(\ell)$. To see this, note that, given $T \geq T^*(\ell)$, the optimal choice of ℓ' is $-\underline{R}$, since $\tilde{V}^{D, \epsilon}(\ell')$ is decreasing in ℓ' by the second part of Lemma E.8. So, the value of the objective function on the right-hand Equation (D) for a given T and $\ell' = \underline{R}$ is

$$-\int_0^T \alpha_h \frac{1 - e^{-(\rho+\lambda)t}}{\rho + \lambda} e^{-rt} dt - e^{-rT} \frac{\int_0^{T^*(\underline{R})} \alpha_h \frac{1 - e^{-(\rho+\lambda)t}}{\rho + \lambda} e^{-rt} dt}{1 - e^{-rT^*(\underline{R})}}$$

But, by Lemma E.8, we know that this is decreasing in T . So among choices $T \geq T^*(\ell)$, it is optimal to set $T = T^*(\ell)$.

I move on to argue the more difficult case when $T \leq T^*(\ell)$. Let $f := \rho + \lambda + r$ and $g := \rho + r$. Apply Lemma E.7 to compute $\tilde{V}^{D, \epsilon}$:

$$\tilde{V}^{D, \epsilon}(\ell) = -\alpha_h \left(\frac{1}{r(\rho + \lambda)} - \frac{1}{(\rho + \lambda)f} \left(1 - e^{-fT^*(\ell)} + e^{-rT^*(\ell)} \left(\frac{1 - e^{-fT^*(\underline{R})}}{1 - e^{-rT^*(\underline{R})}} \right) \right) \right)$$

Now for any T, ℓ' which satisfies the constraint of (*), Lemma E.7 implies that

$$\begin{aligned} & -\int_0^T \alpha_h \frac{1 - e^{-(\rho+\lambda)t}}{\rho + \lambda} e^{-rt} dt + e^{-rT} \tilde{V}^{D, \epsilon}(\ell') \\ &= C_1 + C_2 \left(1 - e^{-fT} + e^{-rT} (1 - e^{-fT^*(\ell')}) + e^{-r(T+T^*(\ell'))} \frac{1 - e^{-fT^*(0)}}{1 - e^{-rT^*(0)}} \right) \end{aligned}$$

where recall $C_2 > 0$. To prove the result, its enough to show that (after canceling):

$$(1 - e^{-fT^*(\ell)}) + e^{-rT^*(\ell)} (1 - e^{-fT^*(\underline{R})}) + e^{-r(T^*(\ell)+T^*(\underline{R}))} \left(\frac{1 - e^{-fT^*(0)}}{1 - e^{-rT^*(0)}} \right)$$

$$\geq \left[1 - e^{-fT} + e^{-rT}(1 - e^{-fT^*(\ell')}) + e^{-r(T+T^*(\ell'))} \frac{1 - e^{-fT^*(0)}}{1 - e^{-rT^*(0)}} \right] \quad (\text{P}^*)$$

for any T, ℓ, ℓ' such that inequality $(*)$ holds, i.e.:

$$x_h \frac{1 - e^{-fT}}{f} - \frac{\rho \bar{R}}{\rho + r} (1 - e^{-gT}) - \ell' e^{-gT} \leq -\ell$$

The next step manipulates this latter inequality into a form that looks like inequality (P^*) . Use the definition of $T^*(\ell')$ to plug in for ℓ' on the left hand side, let $c_A := \frac{\rho \bar{R} - (\rho + r) \underline{R}}{g} f$ and rearrange,

$$x_h(1 - e^{-fT}) + x_h e^{-gT}(1 - e^{-fT^*(\ell')}) - c_A(1 - e^{-g(T+T^*(\ell'))}) + \underline{R}f \leq -\ell f$$

By definition, this holds with equality for $(T, \ell') = (T^*(\ell), -\underline{R})$, so this inequality is equivalent to

$$\begin{aligned} & x_h(1 - e^{-fT}) + x_h e^{-gT}(1 - e^{-fT^*(\ell')}) - c_A(1 - e^{-g(T+T^*(\ell'))}) \\ & \leq x_h(1 - e^{-fT^*(\ell)}) + x_h e^{-gT^*(\ell)}(1 - e^{-fT^*(\underline{R})}) - c_A(1 - e^{-g(T+T^*(\underline{R}))}) \end{aligned}$$

The definition of $T^*(\underline{R})$ implies that $c_A = x_h \frac{1 - e^{-fT^*(\underline{R})}}{1 - e^{-gT^*(\underline{R})}}$. Plugging this in on both sides, rearranging and canceling x_h , we arrive at:

$$\begin{aligned} & 1 - e^{-fT} + e^{-gT}(1 - e^{-fT^*(\ell')}) + e^{-g(T+T^*(\ell'))} \frac{1 - e^{-fT^*(\underline{R})}}{1 - e^{-gT^*(\underline{R})}} \\ & \leq 1 - e^{-fT^*(\ell)} + e^{-gT^*(\ell)}(1 - e^{-fT^*(\underline{R})}) + e^{-g(T^*(\ell)+T^*(\underline{R}))} \frac{1 - e^{-fT^*(\underline{R})}}{1 - e^{-gT^*(\underline{R})}} \end{aligned} \quad (\text{A}^*)$$

Thus, to prove the result, it is sufficient to show that for any T, ℓ', ℓ s.t (A^*) holds, (P^*) holds as well. For any T, ℓ' , let

$$z := e^{-fT^*(\underline{R})}, \quad z_l := e^{-fT^*(\ell)}, \quad u := e^{-fT}, \quad y := e^{-fT^*(\ell')}$$

We have $z \geq z_l$ and $y \leq z$ and by the assumption that $T \leq T^*(\ell)$, we have $u \geq z_l$. Since $T^*(\ell)$ is strictly increasing in ℓ , we have $y = z$ if and only if $u = z_l$. In this case, the deviation is in fact the conjectured optimal choice, and so the inequalities (A^*) and (P^*) hold at equality. Thus, I need only show the result assuming that $u > z_l$ and $y < z$.

Plug these definitions into (A^*) and multiply both sides by $1 - z^{\frac{g}{f}}$ to get

$$\begin{aligned} & (1 - u)(1 - z^{\frac{g}{f}}) + u^{\frac{g}{f}}(1 - y)(1 - z^{\frac{g}{f}}) + (uy)^{\frac{g}{f}}(1 - z) \\ & \leq (1 - z_l)(1 - z^{\frac{g}{f}}) + z_l^{\frac{g}{f}}(1 - z)(1 - z^{\frac{g}{f}}) + (zz_l)^{\frac{g}{f}}(1 - z) \end{aligned}$$

Similarly, plug in to (P^*) and multiply both sides by $1 - z^{\frac{r}{f}}$ to get

$$(1 - u)(1 - z^{\frac{r}{f}}) + u^{\frac{r}{f}}(1 - y)(1 - z^{\frac{r}{f}}) + (uy)^{\frac{r}{f}}(1 - z)$$

$$\leq (1 - z_l)(1 - z^{\frac{r}{f}}) + z_l^{\frac{r}{f}}(1 - z)(1 - z^{\frac{r}{f}}) + (zz_l)^{\frac{r}{f}}(1 - z)$$

Note that in neither case does this multiplication change the direction of the inequality, since $\frac{r}{f}, \frac{g}{f}, z \in (0, 1)$. Each of these is a special case of the inequality, for arbitrary a ,

$$\begin{aligned} & (1 - u)(1 - z^a) + u^a(1 - y)(1 - z^a) + (uy)^a(1 - z) \\ & \leq (1 - z_l)(1 - z^a) + z_l^a(1 - z)(1 - z^a) + (zz_l)^a(1 - z) \end{aligned}$$

After rearranging and canceling terms, we arrive at:

$$\begin{aligned} 0 & \leq (u - z_l) - u^a(1 - y) + (uz)^a(1 - y) - (uy)^a(1 - z) \\ & \quad + z^a(1 - u) + (z_l z^a - z_l^a z) + (z_l^a - z^a) \\ \implies 0 & \leq (u - z_l) + (uz)^a(1 - y) + z_l^a(1 - z) - (uy)^a(1 - z) - u^a(1 - y) - z^a(u - z_l) \end{aligned} \tag{C(a)}$$

The crucial step is the following claim:

if inequality C(\bar{a}) is satisfied at some $\bar{a} \in (0, 1)$, then it is satisfied for all $0 < a \leq \bar{a}$. (C)

The agent's version sets $a = \frac{g}{f}$, while the regulator's version sets $a = \frac{r}{f} < \frac{g}{f}$. Thus, if the agent's version holds – i.e. IC holds – then the regulator's version holds – i.e. value to the deviating choice T, ℓ' is (weakly) lower than the conjectured optimal choice.

To prove this claim, begin by denoting the RHS by $H(a; u, y, z, z_l)$. Instead of showing Property (C) for H , I will show it for $\tilde{H}(a) := \frac{H}{(uy)^a}$, which can be written,

$$\tilde{H}(a) = \frac{(u - z_l)}{(uy)^a} + \left(\frac{z}{y}\right)^a(1 - y) + \left(\frac{z_l}{uy}\right)^a(1 - z) - (1 - z) - \left(\frac{1}{y}\right)^a(1 - y) - \left(\frac{z}{uy}\right)^a(u - z_l)$$

for any $y > 0$, from which Property (C) for H can be recovered immediately. If $y = 0$, the property can be recovered from the right-continuity of H at $y = 0$. There are then, two cases to consider (note that \tilde{H} is smooth in a for any feasible choices of u, y, z, z_l):

Case 1 - $uy < z_l$: How to show this? First, given our baseline ordering of u, z, z_l, y and under the additional assumption that $uy < z_l$, $\tilde{H} \rightarrow \infty$ as $a \rightarrow \infty$. Second, since $uy < z_l$ and $y < z$, $\tilde{H} \rightarrow -(1 - z)$ as $a \rightarrow -\infty$. Third, $\tilde{H}(1) = \tilde{H}(0) = 0$. Given these observations, if I can show that $\frac{\partial \tilde{H}^2}{\partial a^2}$ cross zero at most twice, the result will follow. To see this, observe that to violate the property, there are points $a_0 < 0 < a_1 < a_2 < 1 < a_3$ such that $H(a_0) < 0$, $H(a_1) < 0$, $H(a_2) > 0$, $H(a_3) > 0$, and $H(0) = H(1) = 0$. Connecting these points in an infinitely differentiable way (as required by the definition of H) requires the existence of points $b_0 < b_1 < b_2 < b_3$ such that b_0 and b_2 are local maxima with strictly negative second

derivative while b_1 and b_3 are local minima with strictly positive second derivative. This implies that $\frac{\partial \tilde{H}^2}{\partial a^2}$ crosses zero at least three times. So, as long as I show that $\frac{\partial^2 H(a)}{\partial a^2}$ has at most two zeros, the proof will be complete.

Twice differentiate \tilde{H} ,

$$\begin{aligned} \frac{\partial \tilde{H}^2}{\partial a^2} &= \frac{(u - z_l)}{(uy)^a} \ln\left(\frac{1}{uy}\right)^2 + \left(\frac{z}{y}\right)^a (1 - y) \ln\left(\frac{z}{y}\right)^2 \\ &\quad + \left(\frac{z_l}{uy}\right)^a (1 - z) \ln\left(\frac{z_l}{uy}\right)^2 - \left(\frac{1}{y}\right)^a (1 - y) \ln\left(\frac{1}{y}\right)^2 - \left(\frac{z}{uy}\right)^a (u - z_l) \ln\left(\frac{z}{uy}\right)^2 \end{aligned}$$

We want to show that this object has at most 2 zeros. The zeros of this function are the same as the zeros of the function $G := y^a \frac{\partial \tilde{H}^2}{\partial a^2}$. Computing, we get,

$$\begin{aligned} G &= \frac{(u - z_l)}{(u)^a} \ln\left(\frac{1}{uy}\right)^2 + (z)^a (1 - y) \ln\left(\frac{z}{y}\right)^2 + \left(\frac{z_l}{u}\right)^a (1 - z) \ln\left(\frac{z_l}{uy}\right)^2 \\ &\quad - (1 - y) \ln\left(\frac{1}{y}\right)^2 - \left(\frac{z}{u}\right)^a (u - z_l) \ln\left(\frac{z}{uy}\right)^2 \end{aligned}$$

Then, to show that G has at most two zeros, it would be enough to show that $\frac{\partial G}{\partial a}$ has at most one zero. Differentiating:

$$\begin{aligned} \frac{\partial G}{\partial a} &= \frac{(u - z_l)}{(u)^a} \ln\left(\frac{1}{uy}\right)^2 \ln\left(\frac{1}{u}\right) + (z)^a (1 - y) \ln\left(\frac{z}{y}\right)^2 \ln(z) \\ &\quad + \left(\frac{z_l}{u}\right)^a (1 - z) \ln\left(\frac{z_l}{uy}\right)^2 \ln\left(\frac{z_l}{u}\right) - \ln\left(\frac{z}{u}\right) \left(\frac{z}{u}\right)^a (u - z_l) \ln\left(\frac{z}{uy}\right)^2 \end{aligned}$$

I apply one more transformation: $\frac{\partial G}{\partial a}$ has the same number of zeros as $J := \frac{u^a \frac{\partial G}{\partial a}}{z^a}$. Thus, if $\frac{\partial J}{\partial a}$ is either always negative or always positive, then J has at most one zero and the result will be proved. Differentiating yields

$$\begin{aligned} \frac{\partial J}{\partial a} &= \frac{(u - z_l)}{(z)^a} \ln\left(\frac{1}{uy}\right)^2 \ln\left(\frac{1}{u}\right) \ln\left(\frac{1}{z}\right) + (u)^a (1 - y) \ln\left(\frac{z}{y}\right)^2 \ln(z) \ln(u) \\ &\quad + \left(\frac{z_l}{z}\right)^a (1 - z) \ln\left(\frac{z_l}{uy}\right)^2 \ln\left(\frac{z_l}{u}\right) \ln\left(\frac{z_l}{z}\right) \end{aligned}$$

Recalling that $1 > u > z_l > 0$ and $1 \geq z \geq z_l > 0$ implies that all the terms in the RHS of the above equation are positive and the first is strictly positive. Thus, J is a strictly increasing function and has at most one zero. That implies that the same is true of $\frac{\partial G}{\partial a}$. This therefore implies that G has at most two zeros and hence the same is true of $\frac{\partial \tilde{H}^2}{\partial a^2}$.

Case 2 - $uy \geq z_l$: In this case, I show that for any $a \in (0, 1)$, $H(a) < 0$ (and so the claim is proved).⁴⁵

⁴⁵That is, there can never be a pair (u, y) s.t. $uy \geq z_l$ and (u, y) satisfies IC. Intuitively, it would be as if the regulator said, in between 0 and $T^*(\ell)$, you will have two opportunities for reduced penalties and the

Suppose first that $z_l = uy$. Then, we have:

$$\begin{aligned} H &= (u - z_l) + (uz)^a(1 - y) + z_l^a(1 - z) - (uy)^a(1 - z) - u^a(1 - y) - z^a(u - z_l) \\ &= (1 - y)(u - u^a)(1 - z^a) \end{aligned}$$

and since $u < u^a$ for $a \in (0, 1)$, this is strictly negative. Next, I show that $\frac{\partial H}{\partial z_l} \geq 0$ on $z_l < uy$. This implies that on $z_l \leq uy$, H is maximized at $z_l = uy$, which we've already seen is negative and so the proof will be concluded.

Let $G(z) := \frac{\partial H}{\partial z_l} = z^a + \frac{a}{z_l^{1-a}}(1 - z) - 1$ and differentiate to get $\frac{\partial G}{\partial z} = \frac{a}{z^{1-a}} - \frac{a}{z_l^{1-a}}$. Since $G(1) \geq 0$ and $\frac{\partial G}{\partial z}(z) \leq 0$ for all $z \geq z_l$, we find that $G(z) \geq 0$ for all $z \geq z_l$. This implies by definition that $\frac{\partial H}{\partial z_l}(x) \geq 0$ for all $z \in [z_l, 1]$ and since $z \in [z_l, 1]$ by assumption, this concludes the proof. \square

Proof of Proposition E.1: Suppose that $\theta \in \Theta^*$. Apply Lemma E.9 to conclude that value $V^{D,\epsilon}(\ell)$ for $\epsilon < T^*(0)$ is achieved by the policy functions with $T(\ell) = T^*(\ell)$ and $\ell'(\ell) = \underline{R}$. Observe now that the right-hand side of inequality (DI) in Equation (D) is maximized by setting $R_0 = \underline{R}$, since $V^D(R_0)$ is decreasing in R_0 and R_0 does not otherwise affect the problem. Furthermore, observe that the optimal choice of T_0 is independent of ϵ when $\epsilon < T^*(0)$ is sufficiently small, since the optimal policy in $V^{D,\epsilon}(\ell)$ is independent of ϵ for $\epsilon < T^*(0)$.

So, in summary, for $\epsilon < T^*(0)$, both the optimal policy in $V^{D,\epsilon}$ and the choice T^0 in (DI) are independent of ϵ . This implies that the policy induced by $((T_0, \underline{R}), (T_0 + T^*(\underline{R}), \underline{R}), (T_0 + 2T^*(\underline{R}), \underline{R}), \dots)$ is optimal in problem \mathcal{P}^ϵ and generates a value for the regulator which is independent of ϵ for $\epsilon < T^*(0)$. Applying Lemma E.2 concludes the proof. \square

Proof of Theorem 2: When $\theta \notin \Theta^*$, then an optimal penalty policy sets $R_t = \underline{R}$ for all t . Since the penalty is constant, then as in Proposition 1, it is without loss of generality for the regulator's value to suppose that $A_t(x) = A_s(x)$ for each x and $t, s \geq 0$.

Suppose instead that $\theta \in \Theta^*$. To prove this result, I approximately transform the regulator's problem into one with $\alpha_l = x_l = 0$, to which I will then apply Proposition E.1.

Fix some parameters of the model, $\theta \in \Theta^*$. Rather than studying problem \mathcal{P} , consider the problem

$$V_h^* := \sup_{(R,A) \in \mathcal{M}} V_h(R) \tag{\mathcal{P}^h}$$

where $V_h(R) = \mathbb{E}_{\mathbb{M}(R)} \left[\int_0^\infty N_t^h \alpha_h \right]$. Notably, this problem differs from problem \mathcal{P} only in that losses from the low type agent do not enter the objective function. Let $\tilde{x} := (x_h - x_l)\mathbf{1}_{x_t=x_h}$

second opportunity will have a penalty of 0. By the definition of $T^*(\ell)$, such a policy cannot be incentive compatible.

and $\tilde{R} = \bar{R} - \frac{x_l}{\rho}$. Let $\tau^l := \inf\{t|x_t = x_l\}$ i.e. the transition time into the low state. Then, the agent's value for a given stopping time τ and policy R can be written:

$$\begin{aligned} W(x, t_0, \tau; R) &= \mathbb{E} \left[\int_0^\tau e^{-(\rho+r)t} x_t dt - (1 - e^{-(\rho+r)\tau}) \frac{\rho}{\rho+r} \bar{R} - e^{-(\rho+r)\tau} R_{\tau+t_0} \right] \\ &= \mathbb{E} \left[\int_0^{\tau \wedge \tau_l} e^{-(\rho+r)t} (x_h - x_l) dt + \int_0^\tau e^{-(\rho+r)t} (x_l) dt \right. \\ &\quad \left. - (1 - e^{-(\rho+r)\tau}) \frac{\rho}{\rho+r} \bar{R} - e^{-(\rho+r)\tau} R_{\tau+t_0} \right] \\ &= \mathbb{E} \left[\int_0^\tau e^{-(\rho+r)t} \tilde{x}_t dt - (1 - e^{-(\rho+r)\tau}) \frac{\rho \tilde{R}}{\rho+r} - e^{-(\rho+r)\tau} R_{\tau+t_0} \right] \end{aligned}$$

Notice now that problem \mathcal{P}^h , with the agent's problem written as above, is exactly the problem studied in Proposition E.1, except replacing x_t with \tilde{x}_t . We therefore know that an optimal policy in problem \mathcal{P}^h is (R^*, A^*) defined by:

- $R_{T_0+nT^*} = 0$ for some T_0 with T^* defined by Equation (I)
- $R_t = \tilde{R}$ otherwise
- $A_t^*(x) = 1$ if and only if $t \in \{T_0, nT^*\}_{n \in \mathbb{N}}$

Since $\theta \in \Theta^*$, Theorem 1 implies that $x_l \leq \rho \bar{R} - (\rho+r)\underline{R}$. So, we can apply Lemma 1 to transform (R^*, A^*) into $(\tilde{R}^*, \tilde{A}^*) \in \mathcal{L}$ which has the properties:

- $\tilde{A}_t(x_l) = 1$ for all $t \geq 0$
- $\tilde{A}_t(x_h) = A_t(x_h)$
- $R_t^* = e^{-(\rho+r)(T_0-t)} \underline{R} + (1 - e^{-(\rho+r)(T_0-t)}) \frac{(\rho \bar{R} - x_l)}{\rho+r}$ for $t < T_0$
- $R_t^* = e^{-(\rho+r)(T^* \lceil \frac{t}{T^*} \rceil - t)} \underline{R} + (1 - e^{-(\rho+r)(T^* \lceil \frac{t}{T^*} \rceil - t)}) \frac{(\rho \bar{R} - x_l)}{\rho+r}$ for $t > T_0$ and $t \notin \{T_0 + nT^*\}_{n \in \mathbb{N}}$

where the last two lines translate the last requirement of an element in set \mathcal{L} to the policy R^* . \tilde{R}^* is also an optimal policy for problem \mathcal{P}^h . Lemma 1 further implies that

$$V^* = \sup_{(R,A) \in \mathcal{L}} V(R, A) = V_h^*$$

which concludes proof of the first part of the result.

To conclude the second part of the proof, suppose that $\theta \notin \Theta^*$. Then, by Proposition 1 an optimal penalty policy is $R^{\underline{R}}$, i.e. $R_t = \underline{R}$ for all t , with $A_t(x)$ constant in t for each x . \square

F. Proofs for Appendix A.5

I first present a number of preliminary lemmas. The first three are useful for proving Proposition A.3 and the fourth is necessary for the proof of Proposition A.4. The proofs of the lemmas are tedious and so are relegated to Online Appendix H.

Recall that $\tau^k = \inf\{t|x_t \leq x^k\}$. The first lemma states that if a pure strategy equilibrium exists in which agents arriving in state j report immediately upon transitioning to state x^{j-1} but not before, then this is also an equilibrium for agents arriving in any state $k \geq j$.

Lemma F.1. *Suppose that (τ^{j-1}, τ^{j-1}) is a pure strategy Nash equilibrium for agents arriving at any t_0 with initial state $x_0 = x^j$. Then, (τ^{j-1}, τ^{j-1}) is a pure strategy Nash equilibrium for agents arriving at any t_0 with initial state $x^k \geq x^j$.*

Lemma F.2 states that an agent's value is increasing in his partner's stopping time. This follows immediately from definitions and so the proof is excluded.

Lemma F.2. *Fix any pure strategy τ . Then for any policy (R, p) , arrival time t_0 , initial state x^k , and $\tau_{-i} \leq_{a.s.} \tau'_{-i}$,*

$$W(x^k, t, \tau; \tau_{-i}) \leq W(x^k, t, \tau; \tau'_{-i})$$

The next lemma, Lemma F.3, states that if the minimum of the players stopping times stops with probability 1 by $\hat{\tau}$, then any best-response requires this as well.

Lemma F.3. *Fix any strategy $m = (m_i, m_{-i})$ and stopping time $\hat{\tau}$. Suppose that $R_t < \bar{R}$, $p_{i,t} > 0$ for all t and that*

$$\mathbb{P}_m(\min\{\tau_1, \tau_2\} \leq \hat{\tau}) = 1.$$

Then, if m_i is a best-response to m_{-i} ,

$$\mathbb{P}_m(\tau_i \leq \hat{\tau}) = 1.$$

The following lemma states that if all equilibria for agents arriving in state j at time t_0 involve immediate reporting, then all equilibria for agents arriving at time $t < t_0$ who are in state j at time t_0 must report immediately with probability 1 (or have already reported). This follows from the definition of equilibrium and so its proof is excluded.

Lemma F.4. *Fix a state x^j and some arrival time t_0 . Suppose that $x_0 = x^j$ and every equilibrium, m^* , is such that $\mathbb{P}_{m^*}(\min\{\tau_1, \tau_2\} = 0) = 1$. Then, any equilibrium m^* for arbitrary $x_0 = x^k$ and arrival at $t_0 < t$ is such that $\mathbb{P}_{m^*}(\min\{\tau_1, \tau_2\} > \tau^j) = 0$.*

Now, for any equilibrium m^* , let $\underline{t}_i := \min\{t + t_0 | \mathbb{P}_{m^*}(\tau_i \leq t) = 1\}$ i.e. the first time by which player i stops with probability 1, or ∞ if such a time doesn't exist. Lemma F.5 states

that either agents report with probability 1 immediately, or there is a positive probability of reaching any time, however far in the future. The idea is that if there were some interior time with probability 1 reporting, agents would preempt each other, and this would unravel back to immediate reporting.

Lemma F.5. *Fix any equilibrium m^* , $\epsilon > 0$ and policy $R = R^*(\epsilon, T)$ defined in Section 6. Then,*

$$\underline{t}_i = \{t_0, \infty\}.$$

We now have the lemmas in place necessary to tackle Proposition A.2.

Proof of Proposition A.2: First, I show that $(\tau^{\bar{J}^*}, \tau^{\bar{J}^*})$ is always an equilibrium, independent of (R, p) . By Lemma F.1, it is sufficient to show that $(\tau^{\bar{J}^*}, \tau^{\bar{J}^*})$ is an equilibrium for agents arriving in state $\bar{J}^* + 1$.

To see this is true, suppose that $-i$ uses strategy $\tau^{\bar{J}^*}$. By Lemma F.3, we can restrict to deviations τ such that $\mathbb{P}(\tau \leq \tau^{\bar{J}^*}) = 1$. Then, player i 's net value to any deviation $\tau \leq \tau^{\bar{J}^*}$ is:

$$\begin{aligned} & W(x^{\bar{J}^*+1}, t_0, \tau; \tau^{\bar{J}^*}) - W(x^{\bar{J}^*+1}, t_0, \tau^{\bar{J}^*}; \tau^{\bar{J}^*}) \\ = & \mathbb{E} \left[\left(\int_{\tau \wedge \tau^{\bar{J}^*}}^{\tau^{\bar{J}^*}} (\rho \bar{R} - x_t) e^{-(\rho+r)t} dt \right) + \mathbf{1}_{\tau < \tau^{\bar{J}^*}} \left(e^{-(\rho+r)\tau^{\bar{J}^*}} \frac{\bar{R} + R_{t_0 + \tau^{\bar{J}^*}}}{2} - R_{t_0 + \tau^{\bar{J}^*}} e^{-(\rho+r)\tau} \right) \right] \\ \leq & \mathbb{E} \left[\left(\int_{\tau \wedge \tau^{\bar{J}^*}}^{\tau^{\bar{J}^*}} (\rho \bar{R} - x^{\bar{J}^*+1}) e^{-(\rho+r)t} dt \right) + \mathbf{1}_{\tau < \tau^{\bar{J}^*}} \bar{R} \left(e^{-(\rho+r)\tau^{\bar{J}^*}} - e^{-(\rho+r)\tau} \right) \right] \\ \leq & 0 \end{aligned}$$

where the last inequality follows by the definition of \bar{J}^* . Thus all deviations τ are unprofitable and the conclusion follows.

The next part of the statement is that, given the prescribed policy (R^*, p^*) , for any equilibrium m^* ,

$$\mathbb{P}_{m^*}(\min\{\tau_1, \tau_2\} \leq \tau^{\bar{J}^*}) = 1.$$

If $\bar{J}^* = 0$, this is true since $\tau^0 = \infty$. Otherwise, the proof proceeds by induction on $j \leq \bar{J}^*$. By Lemma F.4, it is enough to show that for an agent arriving in state j at any time, immediate reporting is optimal.

- *Base case, $j = 1$:* In this case, observe that by definition of \bar{J}^* , $x^1 - \rho \bar{R} < 0$ and so if $\mathbb{P}_m(\tau_1 > 0) < 1$, player 2 strictly prefers to immediately report. Therefore, either $\mathbb{P}(\tau_1 = 0) = 1$ and/or $\mathbb{P}(\tau_2 = 0) = 1$. In either case, $\mathbb{P}_m(\min\{\tau_1, \tau_2\} = 0) = 1$.

- *Inductive step, $j \leq \bar{J}^*$* : Apply Lemma F.4 and Lemma F.3 and the inductive hypothesis to see that $\mathbb{P}_{m_1}(\tau_1 \leq \tau^{j-1}) = 1$ in any equilibrium $m = (m_1, m_2)$. Then, player 2's value to any policy, τ , net of his value to immediate reporting, 0, is:

$$\begin{aligned} W(x^j, t_0, \tau; m_1) - W(x^j, t_0, \underline{0}; m_1) &= \mathbb{E} \left[\int_0^{\tau \wedge \tau_1 \wedge \tau^{j-1}} e^{-(\rho+r)t} (x_t - \rho \bar{R}) dt \right. \\ &\quad \left. - e^{-(\rho+r)(\tau \wedge \tau^{j-1})} \mathbf{1}_{\tau \geq \tau^{j-1} \wedge \tau_1} \bar{R} \right] \\ &= C(\tau, m_1)(x^j - (\rho + \lambda) \bar{R}) \end{aligned}$$

where the expectation is taken with respect to m and the distribution of x_t , and where C is some function of τ, m_1 that is strictly negative if $\mathbb{P}_m(\min\{\tau, \tau_1\} > 0) > 0$ and otherwise is 0. By the definition of \bar{J}^* then, $W(x^j, t_0, \tau; \tau_1, R) - W(x^j, \underline{0}, \tau; \tau_1, R)$ is strictly negative and so for any strategy m_1 such that $\mathbb{P}_{m_1}(\tau_1 > 0) < 1$, player 2's best response is immediate stopping. As in the base case, this implies that $\mathbb{P}_m(\min\{\tau_1, \tau_2\} = 0) = 1$. This concludes the inductive step.

Applying Lemma F.4 leads to the conclusion that $\mathbb{P}_m(\min\{\tau_1, \tau_2\} \leq \tau^{\bar{J}}) = 1$. \square

Proof of Proposition A.3: Let,

$$\underline{J}^*(v) := \max\{j | \forall k \leq j, \quad x_j - (\rho + \frac{\lambda}{2}) \bar{R} < -(\rho + r + \frac{\lambda}{2})v\}.$$

I will show that $(\tau^{\underline{J}^*(v)}, \tau^{\underline{J}^*(v)})$ is a pure strategy equilibrium and since $\tau^{\underline{J}^*(0)} \geq \tau^{\underline{J}^*(v)}$ the conclusion follows immediately from this. For any state $j < \underline{J}^*(v) + 1$, the strategy profile amounts to simultaneous immediate reporting which is indeed mutual best-response. For any $j \geq \underline{J}^*(v) + 1$, we can apply Lemma F.1 to see that it is sufficient to show this result for agents with initial state $x_0 = x^{\underline{J}^*(v)+1}$. First, I compute the value to $\tau^{\underline{J}^*(v)}$ given that $-i$ plays $\tau^{\underline{J}^*(v)}$:

$$\begin{aligned} W(x^{\underline{J}^*(v)+1}, t_0, \tau^{\underline{J}^*(v)}; \tau^{\underline{J}^*(v)}, R^v) &= \mathbb{E} \left[\int_0^{\tau^{\underline{J}^*(v)}} (x_t - \rho \bar{R}) e^{-(\rho+r)t} dt - e^{-(\rho+r)\tau^{\underline{J}^*(v)}} \frac{\bar{R} + v}{2} \right] \\ &= \frac{x^{\underline{J}^*(v)+1} - \rho \bar{R} - \frac{\lambda}{2}(\bar{R} + v)}{\rho + r + \lambda} \end{aligned}$$

Any pure deviation τ s.t. $\mathbb{P}(\tau > \tau^{\underline{J}^*(v)}) > 0$ represents a strict loss. Consider instead any pure deviation τ s.t. $\mathbb{P}(\tau < \tau^{\underline{J}^*(v)}) > 0$. In this case, if $W(x^{\underline{J}^*(v)+1}, t_0, \tau; \tau^{\underline{J}^*(v)}) > W(x^{\underline{J}^*(v)+1}, t_0, \tau^{\underline{J}^*(v)}; \tau^{\underline{J}^*(v)})$, there must exist $t \geq t_0$ such that

$$-v > W(x^{\underline{J}^*(v)+1}, t, \tau^{\underline{J}^*(v)}; \tau^{\underline{J}^*(v)}, R^v)$$

where the left-hand side is the value to reporting when $x_t = x^{\underline{J}^*(v)+1}$ and the right-hand side is the value from the strategy $\tau^{\underline{J}^*(v)}$ for an agent arriving in state $x^{\underline{J}^*(v)+1}$ at time t . Plugging in the value on the right-hand side,

$$\begin{aligned} -v &> \frac{x^{\underline{J}^*(v)+1} - \rho\bar{R} - \frac{\lambda}{2}(\bar{R} + v)}{\rho + r + \lambda} \\ \implies -v(\rho + r + \frac{\lambda}{2}) &> x^{\underline{J}^*(v)+1} - (\rho + \frac{\lambda}{2})\bar{R} \end{aligned}$$

But by the definition of $\underline{J}^*(v)$, this inequality cannot hold for any $v \geq 0$. There is thus no pure profitable deviation from $\tau^{\underline{J}^*(v)}$. This implies there is no mixed profitable deviation and I conclude that $(\tau^{\underline{J}^*(v)}, \tau^{\underline{J}^*(v)})$ is a pure strategy Nash equilibrium. \square

Proof of Proposition A.4: I first define a policy $R^*(\epsilon, T)$, choosing ϵ small enough and T large enough for our purposes. Fix $\epsilon > 0$ s.t.

$$\max_{j \leq \bar{J}^*} x^j - (\rho + \lambda^j)\bar{R} + \frac{\epsilon\lambda^j}{2}$$

which is possible by the definition of \bar{J}^* for ϵ sufficiently small. Next fix T such that

$$\frac{1 - e^{-(\rho+r+\lambda^j)T}}{\rho + r + \lambda^j} \left((\rho + \lambda^j)\bar{R} - \frac{\lambda}{2}\epsilon \right) > e^{-fT}x^j$$

for all $j \leq \bar{J}^*$ which is again possible by the definition of \bar{J}^* . Now notice that,

$$-\mathbb{E}_{x^j} \left[\int_0^{T \wedge \tau^{j-1}} (x_t - \rho\bar{R})e^{-(\rho+r)t} dt - \mathbf{1}_{\tau^{j-1} \leq T} \left(\bar{R} - \frac{\epsilon}{2} \right) \right] = \frac{1 - e^{-(\rho+r+\lambda^j)T}}{\rho + r + \lambda^j} \left((\rho + \lambda^j)\bar{R} - \frac{\lambda}{2}\epsilon \right)$$

Then:

$$-\mathbb{E}_{x^j} \left[\int_0^{T \wedge \tau^{j-1}} (x_t - \rho\bar{R})e^{-(\rho+r)t} dt - e^{-(\rho+r)\tau^{j-1}} \mathbf{1}_{\tau^{j-1} \leq T} \left(\bar{R} - \frac{\epsilon}{2} \right) \right] > e^{-fT}x^j \quad (T_\epsilon)$$

where by convention $\tau^0 = \infty$. Let $R = R^*(\epsilon, T)$ where ϵ and T are as defined above. To proceed, fix any t_0 . Then, let

$$t^* := \inf\{t | R_{t_0+t} = 0\}.$$

In words, t^* is the first point after t_0 such that R_{t^*} is 0.

Now, I claim that if $x_0 = x^j$ for $j \leq \bar{J}^*$, then $\mathbb{P}_{m^*}(\min\{\tau_1, \tau_2\} \geq t^*) = 0$. If $\bar{J}^* = 0$, then we are trivially done. Otherwise, I proceed by induction on $j \in \{1, \dots, \bar{J}^*\}$.

- *Base Case, $j = 1$:* Fix some agent i and strategy profile $m = (m_i, m_{-i})$. Consider the

strategy \tilde{m}_{-i} that replaces each τ with

$$\tilde{\tau} := \begin{cases} \tau & \tau \leq t^* \\ \infty & \tau > t^* \end{cases}$$

Then, by Lemma F.2, $W(x^j, t_0, \tau; m_{-i}) \leq W(x^j, t_0, \tau; \tilde{m}_{-i})$ for any τ . For any τ , we also have $W(x^j, t_0, \tau \wedge t^*; \tilde{m}_{-i}) = W(x^j, t_0, \tau \wedge t^*; m_{-i})$ since \tilde{m}_{-i} differs from m_{-i} only after t^* . Thus, for any τ :

$$\begin{aligned} & W(x^j, t_0, \tau \wedge t^*; m_{-i}) - W(x^j, t_0, \tau; m_{-i}) \\ & \geq W(x^j, t_0, \tau \wedge t^*; \tilde{m}_{-i}) - W(x^j, t_0, \tau; \tilde{m}_{-i}) \\ & \geq \mathbb{E} \left[\left(\int_{t^*}^{\tau \wedge (T+t^*)} (\rho \bar{R} - x^1) e^{-(\rho+r)t} dt + e^{-(\rho+r)(T+t^*)} (\bar{R} - \epsilon) \mathbf{1}_{\tau \leq T+t^*} \right. \right. \\ & \quad \left. \left. - e^{-(\rho+r)(T+t^*)} x^1 \mathbf{1}_{\tau > (T+t^*)} \right) \middle| \min\{\tau, \tilde{\tau}_{-i}\} > t^* \right] \times \mathbb{P}_m(\min\{\tau, \tilde{\tau}_{-i}\} > t^*) \\ & \geq \mathbb{E} \left[e^{-(\rho+r)t^*} \left(\int_{t^*}^{T+t^*} (\rho \bar{R} - x^1) e^{-(\rho+r)t} dt - e^{-(\rho+r)T} x^1 \mathbf{1}_{\tau > T} \right) \middle| \min\{\tau, \tilde{\tau}_{-i}\} > t^* \right] \\ & \quad \times \mathbb{P}(\min\{\tau, \tilde{\tau}_{-i}\} > t^*) \\ & \geq C \times \mathbb{P}_m(\min\{\tau, \tilde{\tau}_{-i}\} > t^*) \end{aligned}$$

where C is strictly positive as a result of Equation (T_ϵ) . If, $\mathbb{P}_m(\min\{\tau, \tilde{\tau}_{-i}\} > t^*)$, then we find that

$$W(x^j, t_0, \tau \wedge t^*; m_{-i}) - W(x^j, t_0, \tau; m_{-i}) > 0$$

But, if $\mathbb{P}_m(\min\{\tau, \tilde{\tau}_{-i}\} > t^*) > 0$, it must be that $\mathbb{P}_{m_i}(\tau > t^*) > 0$, a contradiction. The conclusion in this case follows.

- *Inductive Step, $j \leq \bar{J}^*$* : Suppose that the claim is true for all $k < j \leq \bar{J}^*$ and I will show it for j . Apply Lemma F.4 and Lemma F.3 and the inductive hypothesis to see that $\mathbb{P}_{m_1}(\tau_i \leq \tau^{j-1}) = 1$ in any equilibrium $m = (m_1, m_2)$. As in the base case, for any mixed strategy m_{-i} of $-i$, consider the strategy \tilde{m}_{-i} that replaces each τ with $\tau \mathbf{1}_{\tau \leq t^*} + \infty \mathbf{1}_{\tau > t^*}$. Then, because $\mathbb{P}_{m_1}(\tau_i \leq \tau^{j-1}) = 1$, we can apply Lemma F.2 to see that $W(x^j, t_0, \tau_i; m_{-i}) \leq W(x^j, t_0, \tau_i; \tilde{m}_{-i})$. For any τ , we also have $W(x^j, t_0, \tau \wedge t^*; \tilde{m}_{-i}) = W(x^j, t_0, \tau \wedge t^*; m_{-i})$ since \tilde{m}_{-i} differs from m_{-i} only after t^* . Let $\tilde{\tau} := \tau \wedge (t^* + T) \wedge \tau^{j-1}$. For any τ :

$$W(x^j, t_0, \tau \wedge t^*; m_{-i}) - W(x^j, t_0, \tau; m_{-i})$$

$$\begin{aligned}
&\geq W(x^j, t_0, \tau \wedge t^*; \tilde{m}_{-i}) - W(x^j, t_0, \tau; \tilde{m}_{-i}) \\
&\geq \mathbb{E} \left[\left(\int_{t^*}^{\tilde{\tau}} (\rho \bar{R} - x^j) e^{-(\rho+r)t} dt + e^{-(\rho+r)\tau^{j-1}} \mathbf{1}_{\tau^{j-1} \leq \tau \wedge (T+t^*)} \left(\bar{R} - \frac{\epsilon}{2} \right) \right. \right. \\
&\quad \left. \left. + e^{-(\rho+r)\tau} \mathbf{1}_{\tau < \tau^{j-1} \wedge T} \left(\bar{R} - \epsilon \right) - e^{-(\rho+r)T} x_t \mathbf{1}_{\tau \wedge \tau^{j-1} > (T+t^*)} \right) \middle| \min\{\tau, \tilde{\tau}_{-i}\} > t^* \right] \\
&\quad \times \mathbb{P}(\min\{\tau, \tilde{\tau}_{-i}\} > t^*) \\
&\geq \mathbb{E} \left[\left(\int_{t^*}^{T+t^*} (\rho \bar{R} - x_t) e^{-(\rho+r)t} dt + e^{-(\rho+r)\tau^{j-1}} \mathbf{1}_{\tau^{j-1} \leq T+t^*} \left(\bar{R} - \frac{\epsilon}{2} \right) \right. \right. \\
&\quad \left. \left. - e^{-(\rho+r)(T+t^*)} x_t \mathbf{1}_{\tau^{j-1} > T} \right) \middle| \min\{\tau, \tilde{\tau}_{-i}\} > t^* \right] \times \mathbb{P}(\min\{\tau, \tilde{\tau}_{-i}\} > t^*) \\
&\geq C \mathbb{P}(\min\{\tau, \tilde{\tau}_{-i}\} > t^*)
\end{aligned}$$

where C is a strictly positive number as result of Equation (T_ϵ) and the third inequality follows from the definition of \bar{J}^* and ϵ . If, $\mathbb{P}(\min\{\tau, \tilde{\tau}_{-i}\} > t^*)$, these calculations imply that

$$W(x^j, t_0, \tau \wedge t^*; m_{-i}) - W(x^j, t_0, \tau; m_{-i}) > 0$$

But, if $\mathbb{P}(\min\{\tau, \tilde{\tau}_{-i}\} > t^*) > 0$, it must be that $\mathbb{P}(\tau > t^*) > 0$, a contradiction since $\tau \wedge t^*$ is a profitable deviation. It follows that if m is an equilibrium, $\mathbb{P}_m(\min\{\tau_1, \tau_2\} \leq t^*) = 1$. This concludes the inductive step.

To conclude the proof, apply Lemma F.4 and Lemma F.5 (as mentioned at the beginning of the proof) to conclude that for any j and equilibrium m^* , $\mathbb{P}_m^*(\min\{\tau_1, \tau_2\} \geq \tau^{\bar{J}^*}) = 0$. \square

ONLINE APPENDIX

G. Proofs of Lemmas in Appendix E

Proof of Lemma E.1: Observe that for $s \in [T^h(t), t]$, N_s^h is distributed as an arrival-death process with arrival-rate per unit time $\beta = 1$ and death rate per unit time $(\rho + \lambda)N_s^h$, where ρ is the risk of detection while λ is the risk of exit by transition to the low state.

Now, suppose that M_t is an arrival-death process with arrival rate per unit time β and death rate per unit time γM_t .⁴⁶ The expected value of M_t is then

$$\mathbb{E}[M_t] = \beta \frac{1 - e^{-\gamma t}}{\gamma}$$

Applying this formula to our case above leads to the conclusion. \square

Proof of Lemma E.2: Consider any policy $(R, A) \in \mathcal{M}$. If A is such that for all $|t - s| < \epsilon$, $\min\{A_t(x_h), A_s(x_h)\} = 0$, then we are done.

Otherwise, consider the following transformation into a new policy (\tilde{R}, \tilde{A}) :

- Generate \tilde{A} ,
 1. Start with $n = 1$, $t_0 = 0$
 2. Pick any $t_n \in [t_0, t_0 + \epsilon]$ such that $A_{t_n}(x_h) = 0$ and otherwise set $t_n = t_0 + \epsilon$
 3. Let $\tilde{A}_t(x) = \mathbf{1}_{t=t_n}$ for $t \in [t_n, t_n + \epsilon]$
 4. Increment n by 1 and return to step (1).
- If $\tilde{A}_t(x_h) = 0$, let $\tilde{R}_t = \bar{R}$.

Observe that under \tilde{R} , \tilde{A} is an obedient recommendation. Furthermore, under \tilde{A} , the regulator sacrifices value in each interval $[t_n, t_n + \epsilon]$ that converges to 0 as $\epsilon \rightarrow 0$. Because the regulator discounts at a positive rate r , the accumulation of this loss over time converges to 0 as $\epsilon \rightarrow 0$ and the result follows. \square

Proof of Lemma E.3: Let $(T_i, R_i)_{i \leq N}$ be any arbitrary sequence in \mathcal{H}^ϵ . Then let

$$\hat{V}((T_i, R_i)) = - \sum_{i=0}^N e^{-rT_{i-1}} F(T_{i-1}, T_i) - e^{-rT_N} F(T_N, \infty)$$

Then, we want to show that $\hat{V}^* = V^*$, where:

$$\hat{V}^* := \sup_{\mathcal{H}^\epsilon} \hat{V}((T_i, R_i))$$

⁴⁶More precisely, this puts us in the realm of an *immigration*-death model, where the rate of entry of new agents is not affected by the size of the population, while the death rate is a fraction of the population.

First, I will show that $V^* \geq \hat{V}$. Let $(R, A) \in \mathcal{M}^\epsilon$ be any policy in the regulator's problem, \mathcal{P}^ϵ . By Lemma E.1, $\mathbb{E}_{\mathbb{M}(R)} [e^{-rt} N_t] = e^{-rt} \frac{1 - e^{-(t - T^h(t))(\rho + \lambda)}}{\rho + \lambda}$. Let (T_i, R_i) be the sequence of times at which $A_{T_i}(x_h) = 1$ and $R_i = R_{T_i}$. Apply Fubini's theorem to get:

$$\begin{aligned} V(R) &= -\mathbb{E} \left[\int_0^\infty e^{-rt} N_t^h x_h \right] = -\int_0^\infty e^{-rt} \alpha_h \frac{1 - e^{-(\rho + \lambda)(t - T^h(t))}}{\rho + \lambda} \\ &= -\alpha_h \sum_{i \leq N} \int_{T_{i-1}}^{T_i} e^{-rt} \frac{1 - e^{-(\rho + \lambda)(t - T_{i-1})}}{\rho + \lambda} dt - \alpha_h e^{-rT_N} \int_{T_N}^\infty \frac{1 - e^{-(\rho + \lambda)(t - T_N)}}{\rho + \lambda} e^{-r(t - T_N)} dt \\ &= \hat{V}((T_i, R_i)_{i \leq N}) \end{aligned}$$

Thus $V^* \leq \hat{V}^*$. An analogous method shows that $V^* \geq \hat{V}^*$. Let $(T_i, R_i) \in \mathcal{H}^\epsilon$. Then, let $(R^{(T_i, R_i)}, A^{(T_i, R_i)})$ be the induced policy process. From above, we know that $V(R, A) = \hat{V}((T_i, R_i))$ and so we have that $V^* \geq \hat{V}^*$. This concludes the proof. \square

Proof of Lemma E.5: By Lemma E.3 and Lemma E.4, we know that:

$$\begin{aligned} V^{\epsilon, *}&= \sup_{(T_i, R_i)_{i \in N}} -\sum_{i=0}^N e^{-rT_{i-1}} f(T_{i-1}, T_i, \rho + \lambda) dt - e^{-rT_N} f(T_N, \infty, \rho + \lambda) \\ &s.t. \quad C(T_i, T_{i+1}) - e^{-(\rho+r)(T_{i+1}-T_i)} R_{i+1} \leq -R_i \quad \forall i \leq N \\ &\quad T_{i+1} - T_i \geq \epsilon \end{aligned}$$

(with $T_{N+1} = \infty$ if $N < \infty$). Consider now the continuation of the problem after choice of T_0, R_0 , i.e. define:

$$\begin{aligned} V_1^*(a, b) &= \sup_{(T_i, R_i)_{0 \leq i \leq N}} -\sum_{i=1}^N e^{-r(T_{i-1}-T_0)} F(T_{i-1} - T_0, T_i - T_0) - e^{-r(T_N - T_0)} F(T_N - T_0, \infty) \\ &s.t. \quad C(T_i, T_{i+1}) - e^{-(\rho+r)(T_{i+1}-T_i)} R_{i+1} \leq -R_i \quad \forall i \geq 0 \\ &\quad T_{i+1} - T_i \geq \epsilon \quad \forall i \\ &\quad T_0 = a, \quad R_0 = b \end{aligned}$$

Then, we can see that

$$V^{\epsilon, *} = \sup_{R_0 \in [\underline{R}, \bar{R}], T_0 \geq 0} -F(0, T_0) + e^{-rT_0} V_1^*(T_0, R_0)$$

Clearly, $V_1^*(T_0, R_0)$ is independent of T_0 , which implies then that V_1^* can be written unambiguously as a function of just R_0 (abusing notation) and we have in fact:

$$V^{\epsilon, *} = \sup_{R_0 \in [\underline{R}, \bar{R}], T_0 \geq 0} -F(0, T_0) + e^{-rT_0} V_1^*(R_0) \quad (*)$$

I show that $V^D(R_0) \geq V_1^*(R_0)$, which will lead to the conclusion. Fix any sequence $(T_i, R_i)_{i \leq N}$ in the constraint set of $V_1^*(R_0)$.⁴⁷ Then we have,

$$V^{D,\epsilon}(R_0) = -F(0, T_1 - T_0) + V^{D,\epsilon}(R_1)e^{-r(T_1 - T_0)}$$

Proceeding inductively we see that:

$$V^{D,\epsilon}(R_0) = - \sum_{i=0}^J e^{-r(T_i - T_0)} F(0, T_{i+1} - T_i) + e^{-r(T_J - T_0)} V^{D,\epsilon}(R_J)$$

Boundedness of $V^{D,\epsilon}$ implies then that:

$$\begin{aligned} V^{D,\epsilon}(R_0) &= \lim_{J \rightarrow N} - \sum_0^J e^{-r(T_i - T_0)} F(0, T_{i+1} - T_i) + e^{-r(T_J - T_0)} V^{D,\epsilon}(R_J) \\ V^{D,\epsilon}(R_0) &= - \sum_{i=0}^N e^{-r(T_i - T_0)} F(0, T_{i+1} - T_i) \end{aligned}$$

Replacing V_1^* with $V^{D,\epsilon}$ in Equation (*) yields the result. \square

Proof of Lemma E.6: First, observe that $C(0) = 0$. Differentiate $C(T) - e^{-(\rho+r)T} \underline{R}$ with respect to T to get:

$$\frac{\partial (C(T) - e^{-(\rho+r)T} \underline{R})}{\partial T} = e^{-(\rho+r)T} (x_h e^{-\lambda T} - (\rho \bar{R} - (\rho + r) \underline{R}))$$

Since $\theta \in \Theta^*$, $x_h > \rho \bar{R} - (\rho + r) \underline{R}$, $C(T) - e^{-(\rho+r)T}$ is first increasing (starting from \underline{R} at $T = 0$) and then decreasing and converges to $\frac{x_h}{\rho+r+\lambda} - \frac{\rho \bar{R}}{\rho+r}$ — there is thus a unique strictly positive solution to $C(T) - e^{-(\rho+r)T} \underline{R} = -\ell$ for any $\ell \in [\underline{R}, \frac{\rho \bar{R}}{\rho+r} - \frac{x_h}{\rho+r+\lambda}]$ and it is increasing in ℓ . \square

Proof of Lemma E.7: As before, let $f := \rho + r + \lambda$. Plugging in definitions and integrating yields

$$\begin{aligned} & -F(0, T) - e^{-rT} F(0, T') - e^{-r(T+T')} \frac{F(0, T^*(0))}{1 - e^{-rT^*}} \\ &= -\alpha_h \left[\frac{1 - e^{-rT}}{r(\rho + \lambda)} - \frac{1 - e^{-fT}}{f(\rho + \lambda)} + e^{-rT} \left(\frac{1 - e^{-rT'}}{r(\rho + \lambda)} - \frac{1 - e^{-fT'}}{(\rho + \lambda)f} \right) \right. \\ & \quad \left. + e^{-r(T+T')} \left(\frac{1}{r(\rho + \lambda)} - \frac{1 - e^{-fT^*(0)}}{(\rho + \lambda)f(1 - e^{-rT^*(0)})} \right) \right] \\ &= -\alpha_h \left[\frac{1}{r(\rho + \lambda)} - \frac{1}{(\rho + \lambda)f} \left(1 - e^{-fT} + e^{-rT}(1 - e^{-fT'}) + e^{-r(T+T')} \left(\frac{1 - e^{-fT^*(0)}}{1 - e^{-rT^*(0)}} \right) \right) \right]. \end{aligned}$$

⁴⁷It is wlog to suppose that either $N = \infty$ or $T_N = \infty$

Plugging in the definitions of C_1 and C_2 concludes the proof. \square

Proof of Lemma E.8: As before, let $f = \rho + r + \lambda$. Plug T and $T' = T^*(0)$ into Lemma E.7 to get:

$$-F(0, T) - e^{-rT} \frac{F(0, T^*(0))}{1 - e^{-rT^*(0)}} = -\frac{\alpha_h}{r(\rho + \lambda)} + \frac{\alpha_h}{(\rho + \lambda)f} \left(1 - e^{-(\rho+r+\lambda)T} + e^{-rT} \frac{1 - e^{-fT^*(0)}}{1 - e^{-rT^*(0)}} \right)$$

Differentiate the RHS with respect to T to get:

$$\begin{aligned} & \frac{\partial}{\partial T} \left[-F(0, T) - e^{-rT} \frac{F(0, T^*(0))}{1 - e^{-rT^*(0)}} \right] \\ &= \frac{\alpha_h}{\rho + r + \lambda} \left[(\rho + r + \lambda)e^{-fT} - re^{-rT} \frac{1 - e^{-fT^*(0)}}{1 - e^{-rT^*(0)}} \right] \\ &< \frac{\alpha_h}{f} \left[re^{-rT} - re^{-rT} \frac{1 - e^{-fT^*(0)}}{1 - e^{-rT^*(0)}} \right] = \frac{\alpha_h r e^{-rT}}{f} \left[1 - \frac{1 - e^{-fT^*(0)}}{1 - e^{-rT^*(0)}} \right] < 0 \end{aligned}$$

where the third line follows because xe^{-x} is strictly decreasing in x and the last inequality follows because $\frac{1 - e^{-zT}}{1 - e^{-wT}} > 1$ for any $z > w$. To see that $\tilde{V}^D(\ell)$ is increasing ℓ , observe that, by Lemma E.6, $T^*(\ell)$ is increasing in ℓ . Applying the first part of this lemma to $\tilde{V}^D(\ell)$ then leads to the result. \square

H. Proofs of Lemmas in Appendix F

Proof of Lemma F.1: I will show that (τ^{j-1}, τ^{j-1}) is an equilibrium for any t_0 and $x^k \geq x^j$. First, observe that any pure strategy τ can be converted without loss of value to $\tilde{\tau} := \tau \mathbf{1}_{\tau < \tau^j} + \tau^{j-1} \mathbf{1}_{\tau \geq \tau^j}$ by our assumption that (τ^{j-1}, τ^{j-1}) is a pure strategy equilibrium for any t_0 and $x_0 = x^j$. To show that there is no profitable deviation, I need only show that for any T , $\tau^T := \tau^{j-1} \mathbf{1}_{\tau^j \leq T} + T \mathbf{1}_{\tau^j > T}$, because the deviations I need to consider are just mixtures over τ^T and if no τ^T can increase player i 's value, than neither can any mixture. For contradiction, suppose that there exists such a profitable deviation, τ^T .

Observe that, given the assumptions in the lemma and the fact that the generator for x_t is lower bi-diagonal,⁴⁸ there exists m_i s.t.

$$W(x^k, t, m_i; \tau^{j-1}, R) > W(x^j, t, \tau^{j-1}; \tau^{j-1}, R).$$

But then τ^T cannot be strictly preferred to τ^{j-1} since that would imply:

$$-R_{t_0+T} \geq W(x^k, t_0 + T, m_i; \tau^{j-1}, R) > W(x^j, t_0 + T, \tau^{j-1}; \tau^{j-1}, R) > -R_{t_0+T}$$

⁴⁸Lower bi-diagonal because the matrix is 0 except on the main diagonal and the diagonal just below the main diagonal.

a contradiction, where the last inequality follows because (τ^{j-1}, τ^{j-1}) is an equilibrium when agents arrive at $t_0 + T$ in state x^j . Thus, I conclude that in fact τ^{j-1} is weakly optimal among pure strategies τ and thus, among mixtures m_i . \square

Proof of Lemma F.3: Suppose $\mathbb{P}(\min\{\tau_1, \tau_2\} \leq \hat{\tau}) = 1$. Suppose for contradiction that $\mathbb{P}_m(\tau_i \leq \hat{\tau}) < 1$ and m is a best-response. Consider the deviation that replaces every τ in the of support m with $\tau \wedge \hat{\tau}$. As long as $R_t < \bar{R}$ and $p_{i,t} > 0$ for all t , this is a strictly profitable deviation, a contradiction. \square

Proof of Lemma F.5: Suppose not, so that $\underline{t}_{-i} \in (t_0, \infty)$. It is immediate that $\mathbb{P}_{m^*}(\tau_i > \underline{t}_{-i}) > 0$, otherwise i can deviate to $\tau \cap \underline{t}_{-i}$ for a strict improvement (since $\epsilon > 0$). Now, fix some δ small and consider the deviation of player i to $\tau \wedge (\underline{t}_{-i} - \delta)$ for every τ in the support of m_i^* . Denote this deviation by \tilde{m}_i^* . I claim that if $\mathbb{P}_m^*(\tau = \underline{t}_{-i}) > 0$, this is a strictly profitable deviation. To see this, write:

$$\begin{aligned} & W(x^j, t_0, \tilde{m}_i^*; m_{-i}^*, R) - W(x^j, t_0, m_i; m_{-i}^*, R) \\ = & \mathbb{P}(\tau_i = \underline{t}_{-i} \cap \tau_{-i} = \underline{t}_{-i}) \mathbb{E} \left[\int_{\underline{t}_{-i}-\delta}^{\underline{t}_{-i}} e^{-(\rho+r)t} (\rho \bar{R} - x_t) - R_{\underline{t}_{-i}-\delta} e^{-(\rho+r)(\underline{t}_{-i}-\delta)} + \frac{R_{\underline{t}_{-i}} + \bar{R}}{2} e^{-(\rho+r)\underline{t}_{-i}} \right] \\ & + \mathbb{P}(\tau_i = \underline{t}_{-i} \cap \tau_{-i} = \underline{t}_{-i} - \delta) \left[\bar{R} e^{-(\rho+r)(\underline{t}_{-i}-\delta)} - \frac{\bar{R} + R_{\underline{t}_{-i}-\delta}}{2} e^{-(\rho+r)(\underline{t}_{-i}-\delta)} \right] \end{aligned}$$

The second term is weakly positive, while a lower bound on the first term is:

$$\begin{aligned} & \mathbb{P}(\tau_i = \underline{t}_{-i} \cap \tau_{-i} = \underline{t}_{-i}) \mathbb{E} \left[\int_{\underline{t}_{-i}-\delta}^{\underline{t}_{-i}} e^{-(\rho+r)t} (\rho \bar{R} - x_t) - R_{\underline{t}_{-i}-\delta} e^{-(\rho+r)(\underline{t}_{-i}-\delta)} + \frac{R_{\underline{t}_{-i}} + \bar{R}}{2} e^{-(\rho+r)\underline{t}_{-i}} \right] \\ \geq & \mathbb{P}(\tau_i = \underline{t}_{-i} \cap \tau_{-i} = \underline{t}_{-i}) e^{-(\rho+r)(\underline{t}_{-i}-\delta)} \left(-x^J \delta - R_{\underline{t}_{-i}-\delta} + \frac{\bar{R} + R_{\underline{t}_{-i}}}{2} e^{-(\rho+r)\delta} \right) \end{aligned}$$

It can be immediately shown that the term in parentheses on the right-hand side is minimized when $R_{\underline{t}_{-i}-\delta} = \bar{R} - \epsilon$ and $R_{\underline{t}_{-i}} = \bar{R} - \epsilon - \delta$ in which case the term becomes:

$$\left(-x^J \delta - R_{\underline{t}_{-i}-\delta} + \frac{\bar{R} + R_{\underline{t}_{-i}}}{2} e^{-(\rho+r)\delta} \right) = \frac{\bar{R}}{2} e^{-(\rho+r)\delta} + \frac{\bar{R} - \epsilon - \delta}{2} e^{-(\rho+r)\delta} - (\bar{R} - \epsilon) - \Delta x^J$$

which after rearranging and plugging in the assumption that $\epsilon(1 - \frac{1}{2}e^{-(\rho+r)\delta}) > \delta(x^j + e^{-(\rho+r)\Delta}) - \bar{R}(1 - e^{-(\rho+r)\Delta})$, we find to be strictly positive.

Thus, to ensure that m^* is a Nash equilibrium, we must have $\mathbb{P}(\tau_i = \underline{t}_i \cap \tau_{-i} = \underline{t}_{-i}) = 0$. But then, player $-i$ has a strict incentive to use m_{-i}^* s.t. $\mathbb{P}_m^*(\tau_{-i} \leq \underline{t}_{-i} - \delta) = 1$, contradicting

the definition of t_{-i} . The conclusion then follows. \square

I. Proof of Theorem A.1

I prove here that simple Poisson randomized policies do not improve over the deterministic optimal policy.

Proof of Theorem A.1: Let $f := \rho + r + \lambda$. In a γ -Poisson policy, the recommendation $A_t(x_h) = \mathbf{1}_{R_t=0}$ is incentive compatible if and only if

$$\mathbb{E}_{Exp(\gamma)} \left[\frac{1 - e^{-fT}}{f} - \frac{\rho \bar{R}}{\rho + r} (1 - e^{-(\rho+r)T}) \right] \leq 0 \quad (\mathcal{A}^\gamma)$$

The theorem will follow if I can show that any γ -Poisson policy satisfying Equation (\mathcal{A}^γ) is such that

$$\begin{aligned} & \max_{T_0 \geq 0} \int_0^T e^{-rt} \frac{1 - e^{-(\rho+\lambda)t}}{\rho + \lambda} dt \\ & + e^{-rT_0} \frac{-\mathbb{E}_{Exp(\gamma)} \left[\int_0^T e^{-rt} \frac{1 - e^{-(\rho+\lambda)t}}{\rho + \lambda} dt \right]}{1 - \mathbb{E}_{Exp(\gamma)} [1 - e^{-rT}]} \leq \max_{T_0 \geq 0} \int_0^{T_0} e^{-rt} \frac{1 - e^{-(\rho+\lambda)t}}{\rho + \lambda} dt + e^{-rT_0} \frac{-\int_0^{T^*} e^{-rt} \frac{1 - e^{-(\rho+\lambda)t}}{\rho + \lambda} dt}{1 - e^{-rT^*}} \end{aligned}$$

where T^* is defined in Equation (I), the right-hand side is the regulator's value from the optimal deterministic policy, and the left-hand side is the regulator's value from a γ -Poisson policy (and I've canceled α_h). Since the choice of T_0 has the same domain in both problems, it is sufficient to show that

$$\frac{-\mathbb{E}_{Exp(\gamma)} \left[\int_0^T e^{-rt} \frac{1 - e^{-(\rho+\lambda)t}}{\rho + \lambda} dt \right]}{\mathbb{E}_{Exp(\gamma)} [1 - e^{-rT}]} \leq \frac{-\int_0^{T^*} e^{-rt} \frac{1 - e^{-(\rho+r+\lambda)t}}{\rho + \lambda} dt}{1 - e^{-rT^*}} \quad (\mathcal{P}^\gamma)$$

First, I will manipulate \mathcal{A}^γ . From the definition of T^* , we have $\frac{1 - e^{-(\rho+r+\lambda)T^*}}{1 - e^{-(\rho+r)T^*}} = \frac{\rho \bar{R}}{\rho + r} (\rho + r + \lambda)$, which implies

$$\begin{aligned} & \mathbb{E}_{Exp(\gamma)} \left[\frac{1 - e^{-fT}}{f} - \frac{\rho \bar{R}}{\rho + r} (1 - e^{-(\rho+r)T}) \right] \leq 0 \\ \iff & \mathbb{E}_{Exp(\gamma)} \left[1 - e^{-fT} + e^{-(\rho+r)T} \frac{1 - e^{-fT^*}}{1 - e^{-rT^*}} \right] \leq \frac{1 - e^{-fT^*}}{1 - e^{-rT^*}} \quad (\tilde{\mathcal{A}}^\gamma) \end{aligned}$$

Now, I manipulate \mathcal{P}^γ ,

$$\begin{aligned}
& \frac{-\mathbb{E}_{Exp(\gamma)} \left[\int_0^T e^{-rt} \frac{1-e^{-(\rho+\lambda)t}}{\rho+\lambda} dt \right]}{\mathbb{E}_{Exp(\gamma)} [1 - e^{-rT}]} \leq \frac{-\int_0^{T^*} e^{-rt} \frac{1-e^{-(\rho+r+\lambda)t}}{\rho+\lambda} dt}{1 - e^{-rT^*}} \\
& \iff \frac{-\mathbb{E}_{Exp(\gamma)} \left[1 - e^{-rT} - \frac{1-e^{-fT}}{f} \right]}{\mathbb{E}_{Exp(\gamma)} [1 - e^{-rT}]} \leq \frac{-\frac{1-e^{-rT^*}}{r} + \frac{1-e^{-fT^*}}{f}}{1 - e^{-rT^*}} \\
& \iff \frac{\mathbb{E}_{Exp(\gamma)} [1 - e^{-fT}]}{\mathbb{E}_{Exp(\gamma)} [1 - e^{-rT}]} \leq \frac{1 - e^{-fT^*}}{1 - e^{-rT^*}} \\
& \iff \mathbb{E}_{Exp(\gamma)} \left[1 - e^{-fT} + e^{-rT} \frac{1 - e^{-fT^*}}{1 - e^{-rT^*}} \right] \leq \frac{1 - e^{-fT^*}}{1 - e^{-rT^*}} \quad (\tilde{\mathcal{P}}^\gamma)
\end{aligned}$$

Observe that $\tilde{\mathcal{A}}^\gamma$ and $\tilde{\mathcal{P}}^\gamma$ are special cases of the inequality:

$$\mathbb{E}_{Exp(\gamma)} \left[(1 - e^{-fT})(1 - (e^{-fT^*})^a) + (e^{-fT})^a(1 - e^{-fT^*}) \right] \leq (1 - e^{-fT^*}) \quad (C^\gamma(a))$$

where $a = \frac{r}{\rho+r+\lambda}$ for the regulator and $a = \frac{\rho+r}{\rho+r+\lambda}$ for the agent and I've multiplied both sides by $(1 - (e^{-(\rho+r+\lambda)T^*})^a)$. As in Theorem 2, the crucial step is the following:

if $C^\gamma(\bar{a})$ is satisfied at some $\bar{a} \in (0, 1)$, then it is satisfied for all $0 < a \leq \bar{a}$. (C^γ)

With this the proof will be concluded. Let $z = e^{-fT^*}$. Integrating with respect to T , the inequality becomes

$$(1 - z^a)(1 - \frac{\gamma}{\gamma + f}) + \frac{\gamma}{\gamma + fa}(1 - z) - (1 - z) \leq 0$$

Denote the left-hand side by $H(a; z, \gamma)$. Rather than show Property C^γ for H , I will show it for $\tilde{H} = H(\gamma + (\rho + r + \lambda)a)$, from which the property for H can be recovered (since for $a \in [0, 1]$, $sgn(H) = sgn(\tilde{H})$). Computing \tilde{H} , we see that:

$$\tilde{H} = (1 - z^a)(\gamma + fa) - \gamma \frac{\gamma + fa}{\gamma + f}(1 - z^a) + \gamma(1 - z) - (1 - z)(\gamma + fa)$$

I claim that if $\frac{\partial^2 \tilde{H}}{\partial a^2}$ has at most one 0, then Property C^γ will be verified and the proof will be complete. To see this, observe that as $a \downarrow -\infty$, $\tilde{H} \uparrow \infty$. Observe also that $\tilde{H}(0) = \tilde{H}(1) = 0$. To violate the property, there must exist points $a_1 < 0 < a_2 < a_3 < 1$ such that $\tilde{H}(a_1), \tilde{H}(a_2) > 0$, $\tilde{H}(a_3) < 0$, while $\tilde{H}(0) = \tilde{H}(1) = 0$. This requires $\frac{\partial^2 \tilde{H}}{\partial a^2}$ to pass through 0 at least twice. So, I proceed to show that $\frac{\partial^2 \tilde{H}}{\partial a^2}$ has at most one 0.

Twice differentiating \tilde{H} leads to:

$$\frac{\partial^2 \tilde{H}}{\partial a^2} = - \left(\frac{f^2}{\gamma + f} \right) z^a \ln(z) [2 + \ln(z)a]$$

Since $z > 0$, this has exactly one zero at $a = -\frac{2}{\ln(z)}$. So, I conclude that $\frac{\partial^2 \tilde{H}}{\partial a^2}$ crosses 0 at most once so that Property C^γ holds, and the conclusion follows. \square

J. Proof of Proposition A.1

Proof of Proposition A.1: By the assumptions made on x_{hh} and λ_{hh} , the agent never reports in state x_{hh} so that,

$$\mathbb{E}_{\mathbb{M}(R)} \left[\int_0^\infty e^{-rs} \alpha_{hh} \mathbf{1}_{x_{hh},s} \right] = \mathbb{E}_{\mathbb{M}(R')} \left[\int_0^\infty e^{-rs} \alpha_{hh} \mathbf{1}_{x_{hh},s} \right]$$

for any R, R' where C is the constant loss from the x_{hh} types which is independent of the policy. Then

$$V_\gamma(R) := C - \mathbb{E}_{\mathbb{M}(R)} \left[\int_0^\infty e^{-rs} (\alpha_h \mathbf{1}_{x_h,s} + \alpha_l \mathbf{1}_{x_l,s}) ds \right]$$

The regulator's problem can be viewed as one in which the agent arrives at Poisson at rate γ from 0, which the regulator does not observe.

Let $f := \rho + r + \lambda$. Suppose first that $x_l = \alpha_l = 0$. The proofs of Lemmas E.3 to E.6 given in Appendix E are unchanged, except replacing $F(T, T')$, V^* and \tilde{V}^D with $F_\gamma(T, T')$, V_γ^* and \tilde{V}_γ^D , respectively, defined below.

$$\begin{aligned} F_\gamma(T, T') &= \int_T^{T'} \gamma e^{-\gamma(t-T)} \int_t^{T'} e^{-r(s-T)} e^{-(\rho+\lambda)(s-t)} ds dt \int_0^{T'-T} \gamma e^{-\gamma t} \int_t^{T'-T} e^{-rs} e^{-(\rho+\lambda)(s-t)} ds dt \\ &= \int_0^{T'-T} \gamma e^{-(\gamma-(\rho+\lambda))t} \int_t^{T'-T} e^{-rs} e^{-(\rho+\lambda)s} ds dt = \int_0^{T'-T} \gamma e^{-(\gamma+r)t} \frac{e^{-ft} - e^{-f(T'-T)}}{f} dt \\ &= \frac{\gamma}{f} \left[\left(\frac{1}{\gamma+r} + \frac{1}{\rho+\lambda-\gamma} \right) (1 - e^{-(\gamma+r)(T_i-T_{i-1})}) - \frac{1 - e^{-f(T_i-T_{i-1})}}{\rho+\lambda-\gamma} \right], \end{aligned}$$

$$V_\gamma^* = \sup_{\mathcal{H}} - \sum_{i=0}^N e^{-(r+\gamma)T_{i-1}} F_\gamma(T_{i-1}, T_i) - e^{-(r+\gamma)T_N} F_\gamma(T_N, \infty)$$

$$\tilde{V}_\gamma^D(\ell) = F(0, T^*(\ell)) + e^{-(r+\gamma)T^*(\ell)} \frac{F(0, T^*(0))}{1 - e^{-(r+\gamma)T^*(0)}}$$

where, as in the proof of Theorem 2, \mathcal{H} denotes the set of all sequences $(T_i, R_i)_{i \leq N}$ that

induce obedient recommendations. An adjusted Lemma E.7 holds as well for any T, T' :

$$\begin{aligned}
& -F_\gamma(0, T) - e^{-rT}F_\gamma(0, T') - e^{-(r+\gamma)(T+T')} \frac{F_\gamma(0, T^*(0))}{1 - e^{-(r+\gamma)T^*}} \\
= & -\frac{\gamma}{f} \left[\left(\frac{1}{\gamma+r} + \frac{1}{\rho+\lambda-\gamma} \right) (1 - e^{-(\gamma+r)T}) - \frac{1 - e^{-fT}}{\rho+\lambda-\gamma} \right] \\
& - e^{-(r+\gamma)T} \left[\frac{\gamma}{f} \left[\left(\frac{1}{\gamma+r} + \frac{1}{\rho+\lambda-\gamma} \right) (1 - e^{-(\gamma+r)T'}) - \frac{1 - e^{-fT'}}{\rho+\lambda-\gamma} \right] \right] \\
& - e^{-(r+\gamma)(T+T')} \left[\frac{\gamma}{f} \left[\left(\frac{1}{\gamma+r} + \frac{1}{\rho+\lambda-\gamma} \right) (1 - e^{-(\gamma+r)T^*(0)}) - \frac{1 - e^{-(\rho+r+\lambda)T^*(0)}}{\rho+\lambda-\gamma} \right] \right] \\
& \qquad \qquad \qquad \frac{1 - e^{-(\gamma+r)T^*(0)}}{1 - e^{-(\gamma+r)T^*(0)}} \\
= & -C_1(1 - e^{-(\gamma+r)T}) + C_2(1 - e^{-fT})s + e^{-(\gamma+r)T} \left(-C_1(1 - e^{-(\gamma+r)T'}) + C_2(1 - e^{-fT'}) \right) \\
& + e^{-(\gamma+r)(T+T')} \left(\frac{-C_1(1 - e^{-(\gamma+r)T^*(0)}) + C_2(1 - e^{-(\rho+r+\lambda)T^*(0)})}{1 - e^{-(\gamma+r)T^*(0)}} \right) \\
= & -C_1 + C_2 \left[1 - e^{-fT} + e^{-(r+\gamma)T}(1 - e^{-(\rho+r+\lambda)T}) + e^{-(r+\gamma)(T+T')} \frac{1 - e^{-fT^*(0)}}{1 - e^{-(\gamma+r)T^*(0)}} \right]
\end{aligned}$$

with $0 < C_1, C_2$. To see that an adjusted Lemma E.8 holds as well, we need

$$-F_\gamma(0, T) - e^{-(r+\gamma)T} \frac{F_\gamma(0, T^*(0))}{1 - e^{-(\gamma+r)T^*(0)}}$$

to be decreasing in T for any $T \geq T^*(0)$. But, under the assumption that $\gamma < \rho$, the exact same steps can be taken to show this.

Lemma E.9 proceeds in exactly the same (the only difference is that now the regulator applies an extra discounting, γ , across (but not within) periods of time in which no agents report). As long as $\gamma < \rho$, the exact same proof works, but with F_γ replacing F . The remaining steps follow as in Appendix E. \square