Should I Stalk or Should I Go?
An Insurance Auditing Exploration/Exploitation Dilemma

February 15, 2019

Abstract

We consider an insurance fraud problem where policyholders and service providers may collude. When interactions are repeated between the auditor (insurer) and auditees (service providers), auditing behaves as an information gathering mechanism to separate the wheat (honest agents) from the chaff (defrauders). We analyze a Bayesian insurer's dynamic auditing problem under ex-post moral hazard, and characterize his optimal strategy as a strategic exploration/one-armed bandit one. We show that the insurer faces the well-known reinforcement learning exploration/exploitation trade-off between gathering information for higher future profits (exploration) and prioritizing immediate profits (exploitation). We characterize optimal auditing strategies and how exploration levels vary across time. We also show how the degree of informativeness of auditing (i.e., its separating power) and the insurer's time preference affect optimal strategies.

Keywords: insurance fraud; dynamic optimal auditing; strategic exploration; one-armed bandit

JEL: D82; D83; G22; L21

1 Introduction

When the principal can monitor the agent at some cost, principal-agent problems come down to a balancing between investigation-generated gains and costs. Eventually, for strategic agents, principal's optimal strategies may also rely on monitoring as a deterrence device. This paper focuses on the case when interactions are repeated and the principal can acquire time-invariant[1] information through his monitoring efforts. Can the principal be mistaken at first, with who was initially believed to be a "bad-type" agent revealing to actually be a "good-type"? Could it be beneficial for the principal to take some time to learn more about a "bad" agent who might be a "good" one, at the risk of not changing his beliefs afterwards? If so, can we characterize an optimal discovering process?

This problem appears in many situations such as tax auditing and monitoring by fair-trade authorities. It also arises in insurance markets prone to ex-post moral hazard in the form of insurance claim fraud. The insurer (principal) interacts regularly with Service Providers[2] (agents). This repeated aspect is even more relevant in health insurance, because collusion with a Service Provider (hereafter SP) is a necessary condition[3] for fraud to take place, and it may even happen that an SP defrauds alone by submitting a false claim in the name of a policyholder.[4] The relevance of repeated interactions in the SP case is justified by the fact that a single SP may treat several policyholders of the same insurer, and will thus interact with the insurer every time one of them has an accident, which means more frequently than a single policyholder. In addition, in health insurance, SPs have strong partnerships with insurers through the existence of affiliation networks,[5] and insurers can leverage it and threaten with disaffiliation in case of fraud.

This repeated character of interactions is the central feature of the economic phenomenon under scrutiny, because the information gathered through an audit only becomes useful at subsequent periods. And the remaining interaction opportunities also matter because this information is time-invariant and can be reused for all future audits. A preliminary analysis of the role of information gathering in a dynamic auditing setting is provided by Aboutajdine & Picard (2018): in a simple two-period model, we prove that the insurer will optimally audit some unprofitable claims at the first period, because the information acquired allows for a more profitable audit at the last period. On the other hand, optimal auditing at the second period is myopic and excludes unprofitable claims, due to any information acquisition being useless since there are no more future interactions left. In this paper, we extend this setting to an arbitrary number of periods and confirm that information gathering is encouraged by the existence of future interaction opportunities to make use of this information, with more information gathering efforts made early in the relationship. We show that the insurer's problem is a strategic experimentation one, where he faces an exploration/exploitation dilemma between auditing or not auditing given his current information on immediate payoffs (exploitation), and systematically auditing to refine his beliefs, which will only be useful in the future (exploration).

[1] e.g., agent types.

[2] Mechanics, doctors, opticians, ...

[3] e.g., a doctor's certification that the patient is ill is mandatory

[4] Some systems allow the insurer to directly reimburse a practitioner without the policyholder as an intermediary. See for example the Tiers-Payant system in France.

[5] Preferred Provider Organizations (PPOs) in the US, Réseaux de Soins in France

In this context, it is equivalent to playing on a one-armed bandit[6] machine with a risky arm which may yield positive or negative payoffs (auditing), and a safe arm with payoff 0 (no auditing). We then define measures of the degree of informativeness of auditing and characterize its impact and that of time discounting on the insurer's propensity to explore. Unsurprisingly, we find that the more informative an audit, the more exploration takes place, and that the more patient the insurer, the more exploration takes place.

This informational role of auditing is to be compared to the two traditional and well-identified roles of auditing. The first role is to recover undue claims[7] and is actually embedded in our model. It corresponds to a pure exploitation situation when the insurer is myopic[8]. The second role is that of a deterrence mechanism: because an agent's propensity to defraud decreases with the probability of being audited, an increase in the monitoring efforts proves to be beneficial, and Dionne et al. (2008) show that it is optimal to go beyond the myopic optimum and audit some individually unprofitable claims. This deviation from the myopic strategy is, at first sight, very similar to the deviation resulting from the informational role of auditing, but the motivations are very different and the amplitude of the deterrence deviation is time-independent. The fashion in which learning takes place in our setting can also be interpreted in terms of (in)adequate learning situations as defined in Aghion et al. (1991), in which the authors characterize the conditions under which different levels of information gathering should be reached. We also exhibit common features between our model under maximal informativeness[9] and to Rosenberg et al. (2007)'s model of one-armed bandit problems with social learning when the number of players tends to infinity. Our model also provides a new rationalization of the practice of early auditing for which a common explanation is an education effect. For example, in an experimental setting, Guala & Mittone (2005) identify what they call an "echo effect": auditees subject to investigations early in the relationship are less prone to defraud later on, while those who are audited later are less likely to be deterred once they get used to defrauding. From our article's point of view, early auditing is motivated by the numerous opportunities (i.e., remaining periods) to use the acquired information later on.

The rest of the paper is organized as follows. In Section 2, we begin by introducing the model, reformulating the problem as a dynamic programming one and providing an intuitive illustration of the role of auditing as a separating mechanism. In Section 3, we prove our main results, firstly that the auditing set of targeted auditees becomes larger as the number of remaining periods increases, and, secondly, that there is a maximal amount of exploration asymptotically, where the limit corresponds to the infinite-horizon time-independent solution. In Section 4, we focus on the impact of the degree of informativeness and time discounting on the asymptotic optimal auditing sets. Section 6 concludes. Proofs are in the Appendix.

[6] A one-armed bandit has actually two arms.

[7] see Picard (2013) for a comprehensive review

[8] Or, equivalently, with a one-period model

[9] i.e., a single audit perfectly reveals types

2 The Model

2.1 Setting

We consider a world inhabited by an auditor (the insurer), and a continuum (mass 1) of auditees (the service providers, denoted SPs). SPs play a fundamental role in that they act as intermediaries between the insurer and his policyholders, and since an SP usually serves several policyholders, he frequently interacts with the insurer. He intervenes as a trusted third party by certifying the veracity of the claims and his participation is a necessary condition for fraud to occur. We therefore restrict the potential auditees to the population of SPs and leave aside the policyholders. Each SP channels one and only one claim at each period with value 1, each claim being either valid or invalid. The invalidity of a claim may result from honest mistakes or actual ill-disposed voluntary intent to defraud. The probability of submitting invalid claims depends on the SP's intrinsic type, with each SP being either honest (type H, transmit invalid claims only involuntarily) or dishonest (type D, transmit invalid claims both voluntarily and involuntarily). Because of this, D types are more prone to submitting invalid claims than H types, and we consider that they do so with probability p_D, while this probability is p_H for H types, with $p_H < p_D$.

We generalize the two-period model considered by Aboutajdine & Picard (2018) to the case of an arbitrary number of periods T. Each period is indexed by $t \in \{1, 2, .., T\}$ and the initial period $t = 1$ is the beginning of the insurer-SP relationship. In particular, the total number of periods over which the interactions take place may be interpreted as the "lifespan" of an SP. For example, if an insurer checks the channeled claims every month and a pharmacist works for 30 years (i.e., 360 months), then $T = 360$. Of course, the beginning and end of the interactions between the insurer and the different SPs are not the same in reality, but for the sake of simplicity, we consider that all SPs start and finish their relationship with the insurer at the same time. The time-dependent variables of interest will be indexed by both the period at which they are considered and the total number of periods, i.e., by $(t, T) \in \{0, 1, .., T\} \times \mathbb{N}$.

Types are of course time-invariant private information, unobservable to the insurer. But the latter, based on his experience and his observations, has a subjective belief $\pi_{t,T} = \mathbb{P}(D|\pi_{t,T})$ that a given SP is of type D. This prior $\pi_{t,T}$ is distributed in $(0, 1)$ with density $f_{t,T}(\pi)$ and c.d.f $F_{t,T}(\pi)$. The period 0 prior may be considered as initialized at some arbitrary value (e.g., 0.5 if no relevant information is available) or based on some other observable characteristics of the SP at the start of the relationship. An SP to whom the auditor assigns a prior $\pi_{t,T}$ will therefore, from the insurer's point of view, submit an invalid claim with probability $\bar{p}(\pi_{t,T}) = (1 - \pi_{t,T})p_H + \pi_{t,T}p_D$. The insurer may audit any claim and reveal its true status, valid or invalid. This decision is made at each period and is represented by the choice of an auditing strategy $x_{t,T} \in [0, 1]$, which is the probability with which an auditor will investigate the claim submitted by a given SP. Since the only information available to the auditor is the prior, the auditing strategy is a function of $\pi_{t,T}$

$$\begin{aligned} x_{t,T}(\cdot) : [0,1] &\longrightarrow [0,1] \\ \pi_{t,T} &\longmapsto x_{t,T}(\pi_{t,T}). \end{aligned}$$

It costs $c \in (p_H, p_D)$ to perform an audit, inducing a net proceed of $1 - c$ when a claim is found invalid, and $-c$ when it is found valid. In expectation, the net proceed of auditing an SP with prior $\pi_{t,T}$ is $\bar{p}(\pi_{t,T}) - c$. In particular, if the type was known, auditing a dishonest SP would yield an expected net proceed of $p_D - c > 0$ while auditing an honest one would yield $p_H - c < 0$. This difference in the profitability of auditing between types of SPs is at the heart of the insurer's decision on whether to perform one. It is related to the aforementioned first role of auditing, that is to get back the illegitimate reimbursements.

Period t audits also allow the insurer to update his beliefs at the beginning of period $t + 1$. Depending on whether an audit has been performed and, if so, whether the claim was valid or invalid (*Val* and *Inv*, respectively), posterior beliefs $\widetilde{\pi}_{t+1,T}$ are deduced from initial beliefs $\pi_{t,T}$ through Bayes' Law:

$$\widetilde{\pi}_{t+1,T} = \begin{cases} \mathbb{P}(D|audit, Inv) = A(\pi_{t,T}) = \frac{p_D \pi_{t,T}}{\bar{p}(\pi_{t,T})}, \\ \\ \mathbb{P}(D|audit, Val) = B(\pi_{t,T}) = \frac{(1-p_D)\pi_{t,T}}{1-\bar{p}(\pi_{t,T})}, \\ \\ \mathbb{P}(D|no\ audit) = \pi_{t,T}, \end{cases} \tag{1}$$

with

$$\begin{aligned} &B(\pi_{t,T}) < \pi_{t,T} < A(\pi_{t,T}), \\ &A' > 0, \quad A'' < 0, \\ &B' > 0, \quad B'' > 0. \end{aligned}$$

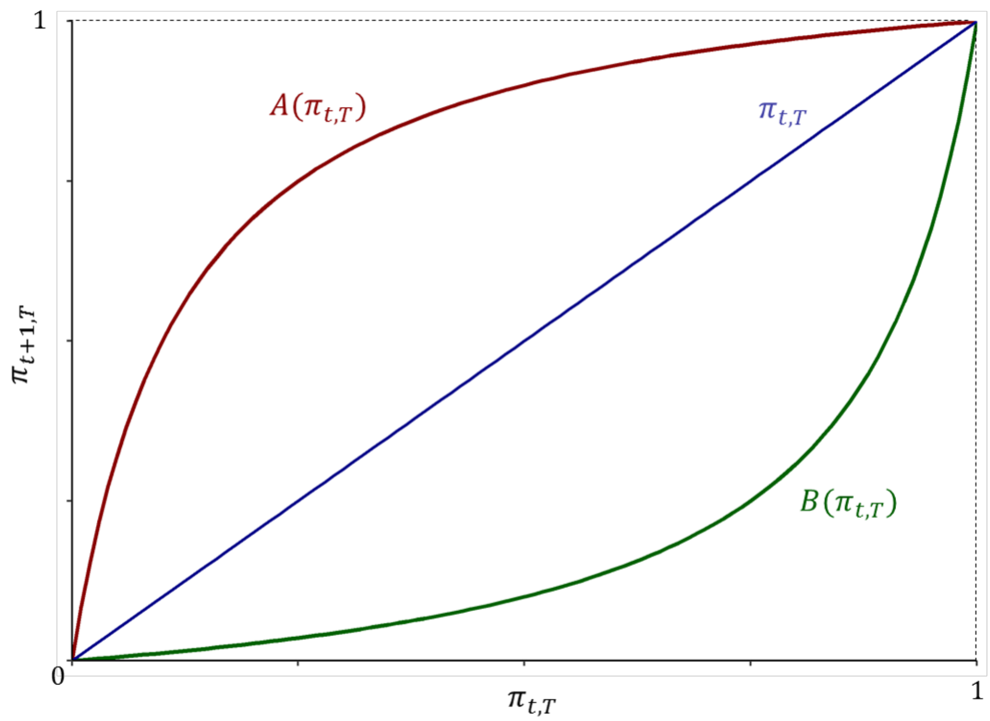


Figure 1: Updating Priors According to the Auditing Status ($p_D = 0.9, p_H = 0.1$)

The updating process is such that an invalid claim increases the belief that an SP is dishonest ($\pi_{t+1,T} = A(\pi_{t,T}) > \pi_{t,T}$) while a valid claim decreases it ($\pi_{t+1,T} = B(\pi_{t,T}) < \pi_{t,T}$). Either way, both cases imply that the insurer acquires information about the auditee's true type. On the contrary, not auditing leaves the beliefs unchanged since no relevant information is obtained. This is akin to the second role of auditing, the acquisition of information.

2.2 One-period optimal auditing

When $T = 1$, we are in presence of the original naive myopic model of auditing, where the investigator only targets individually profitable claims.

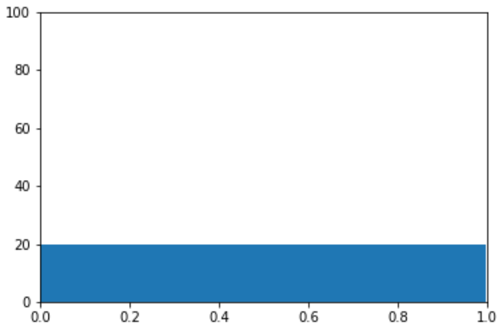
Lemma 1. *For $T = 0$, the optimal auditing strategy is characterized by*

$$\pi^+ \quad s.t. \quad \bar{p}(\pi^+) - c > 0.$$

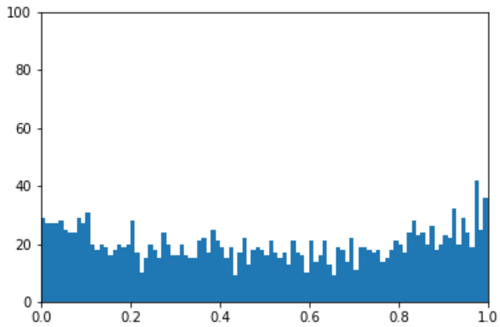
This is a pure exploitation strategy, which is consistent with the fact that there is no subsequent interaction, so there will be no opportunity to use the gathered information.

2.3 Pure Exploration: an intuitive illustration

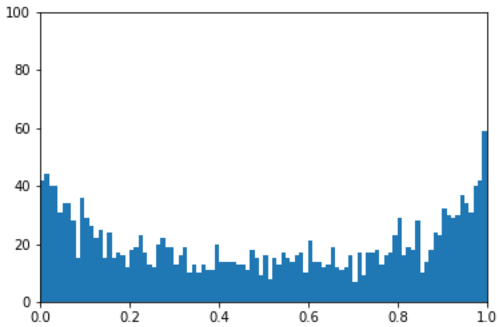
An important feature of bayesian updating is a constantly revised prior end up converging to the true probability it represents. In our case, that would mean that an SP of type D who is constantly audited would see the associated prior converge to 1, and a type H prior would converge to 0. In other words, auditing is a separating tool between dishonest SPs and honest ones.



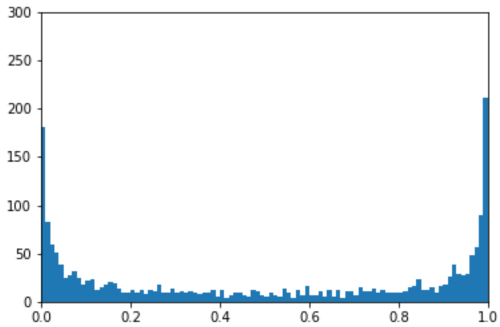
(a) $t = 0, T = 999$



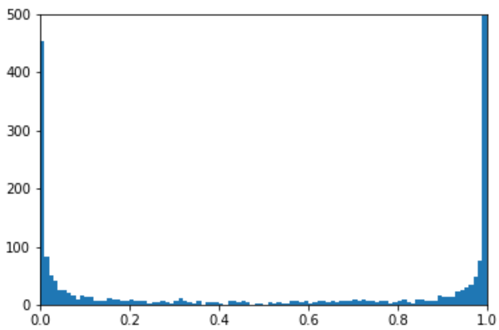
(b) $t = 50, T = 999$



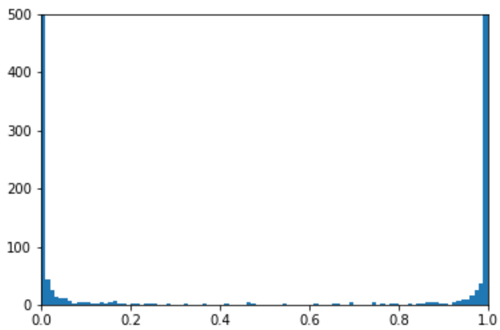
(c) $t = 100, T = 999$



(d) $t = 250, T = 999$



(e) $t = 500, T = 999$



(f) $t = 999, T = 999$

Figure 2: Distribution $f_{t,T}(\pi)$ of priors with repeated audits

Figure 2 illustrates the convergence of priors to both ends of the spectrum. After enough repeated auditing, dishonest SPs are distinguishable from honest ones with a high degree of certainty. At this point, the insurer can select the higher priors and get higher proceeds. But, doing so, he also pays the price of auditing honest SPs at the earlier periods to identify them. Such a strategy where auditing is performed systematically corresponds to a pure exploration approach to the problem.

Another alternative is to restrict auditing to the population of priors for which the expected proceeds of auditing are positive. In this case, the auditor does not pay as much investigation costs for SPs who will end up being honest ones, but some dishonest SPs who started with low priors may never be identified as so. This last approach is a pure exploitation one.

It happens that the optimal auditing thresholds offer a measure of the balance between pure exploration and pure exploitation. Let π^+ be the threshold verifying

$$\bar{p}(\pi^+) - c = 0, \quad \pi^+ = \frac{c - p_H}{p_D - p_H}.$$

The prior π^+ is the lower prior such that expected proceeds of auditing are positive. Therefore, it is a upper bound for optimal auditing thresholds and the closer a threshold to π^+ the more exploitation. The closer a threshold to 0, the more exploration takes place. We can therefore identify π^* as a measure of the balance between exploitation and exploration.

Figure 3: Levels of Exploration/Exploitation

Another phrasing of our goal hereafter is to understand what is the optimal balance between the pure exploration and the pure exploitation strategies.

2.4 Dynamic Programming Optimization Problem

2.4.1 Optimal Auditing Strategies

Let us consider the decision problem faced by the insurer. At any time $t \in \{1, .., T\}$, he must choose an auditing strategy $x_{t,T}(.) : \pi \in [0,1] \to [0,1]$ defining the probability of monitoring SPs for each prior π in order to maximize the corresponding expected proceeds of auditing. For a given $T \in \mathbb{N}$, for any $t \in \{1, .., T\}$, let $\mathcal{V}_{t,T}$ be the expected proceeds from auditing at period t, cumulated over periods $\{t, t+1, .., T\}$, discounted at time t,

$$\mathcal{V}_{t,T} = \max_{x_{t,T}(\cdot)} \left\{ \int_0^1 \Big(\bar{p}(\pi) - c\Big) x_{t,T}(\pi) f_{t,T}(\pi) d\pi + \delta \mathbb{E}_t \Big[\mathcal{V}_{t+1,T} | x_{t,T}(\cdot) \Big] \right\},$$
$$s.t. \quad 0 \le x_{t,T}(\pi) \le 1 \quad \forall \pi \in (0,1),$$

and let the sequence $\left\{x^*_{t,T}(\cdot)\right\}_{t \in \{0,..,T\}}$ denote the corresponding optimal solution.

The objective function can be decomposed into two components. The first component, $\int_0^1 \Big(\bar{p}(\pi) - c\Big) x_{t,T}(\pi) f_{t,T}(\pi) d\pi$, represents the current expected proceeds of auditing when choosing strategy $x_{t,T}(\cdot)$. It is related to the role of auditing as a recovering tool. As for the second component, $\mathbb{E}_t \Big[\mathcal{V}_{t+1,T} | x_{t,T}(\cdot) \Big]$, it represents the maximized (at period $t+1$) future accumulated proceeds of auditing given the choice of $x_{t,T}(\cdot)$ at period t. This conditioning implies that the auditor's choices in the present affects future proceeds, which may not seem intuitive at first. This is because, while the impact of the current strategy on current proceeds is straightforward, its influence on future ones is more indirect.

As it alters the objective function, the inclusion of this second component may modify the optimal solution, depending on how it behaves when the choice of $x_{t,T}(\cdot)$ changes. Hence, we need to better understand how each component is affected by auditing decisions at first, then to identify how these two roles of auditing may be either conflicting or complementary. Obviously, when both components are affected in the same direction, the problem is simple. When these directions are different, we must characterize their amplitude to understand which effect dominates. However, the optimization problem as presented above corresponds to a maximization program in an infinite-dimensional functional space, which makes it difficultly tractable. In order to be able to answer the questions of interest, the problem must be transformed.

2.4.2 Pointwise Maximization Approach

A simplifying reformulation is indeed possible, noticing that the problem at stake has a pointwise maximization structure and can be tackled by considering each prior separately. In this case, the insurer can set an auditing probability $x_{t,T}$ for each prior $\pi_{t,T} \in (0,1)$ at any period t by solving

$$
\begin{aligned}
V_{t,T}(\pi_{t,T}) &= \max_{x_{t,T} \in [0,1]} \left\{ \Big[\bar{p}(\pi_{t,T}) - c\Big] x_{t,T} + \delta \sum_{\pi' \in U(\{\pi_{t,T}\})} \mathbb{P}(\pi' | \pi_{t,T}, x_{t,T}) V_{t+1,T}(\pi') \right\} \\
&= \max_{x_{t,T} \in [0,1]} \left\{ \Omega_{t,T}(\pi_{t,T}, x_{t,T}) \right\},
\end{aligned}
$$

where

$$
U(\{\pi\}) = \{B(\pi), \pi, A(\pi)\}.
$$

This is possible because there is no constraint that may imply a trade-off between auditing intensities for different priors.[10]

The quantity $V_{t,T}(\pi)$ is the cumulated proceeds of auditing an SP with prior π at time t during periods $t, t+1, .., T$. It verifies, for all $t \in \{1, .., T\}$, that

$$
V_{t,T}(\pi_{t,T}) = \Omega_{t,T}(\pi_{t,T}, x^*_{t,T}(\pi_{t,T})),
$$

and

$$
\mathcal{V}_{t,T} = \int_0^1 V_{t,T}(\pi) f_{t,T}(\pi) d\pi.
$$

[10]For example, a constrained auditing budget C, i.e., $\int_0^1 x(\pi) c d\pi \leq C$, would change the pointwise structure of the problem.

3 Exploration vs Exploitation: auditing to separate the wheat from the chaff

3.1 Indicator Optimal Strategies and Optimal Auditing Thresholds

We will now characterize the form of the optimal auditing strategies, for which we will consider the F.O.C at time t. This condition is very simple due to the linearity in $x_{t,T}$ of the objective function

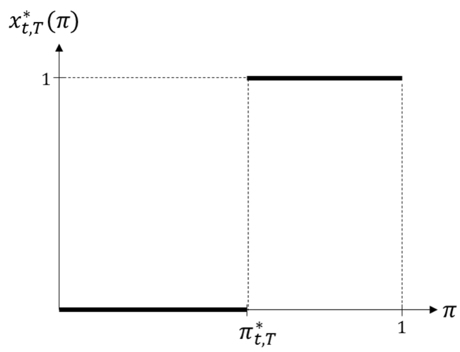
$$\frac{\partial \Omega_{t,T}}{\partial x_{t,T}}\Big|_{\pi_{t,T}} = \bar{p}(\pi_{t,T}) - c + \delta\Big[\bar{p}(\pi_{t,T})V_{t+1,T}(A(\pi_{t,T})) + (1-\bar{p}(\pi_{t,T}))V_{t+1,T}(B(\pi_{t,T})) - V_{t+1,T}(\pi_{t,T})\Big].$$

The F.O.C is independent of $x_{t,T}$, which means we will obtain corner solutions $x^*_{t,T}(\pi_{t,T}) \in \{0,1\}$, depending on the sign of the derivative. It implies that the optimal strategies take a very specific form, such that an SP is either audited with certainty or not audited at all, depending on his prior. This is formalized by Lemma 2 below.

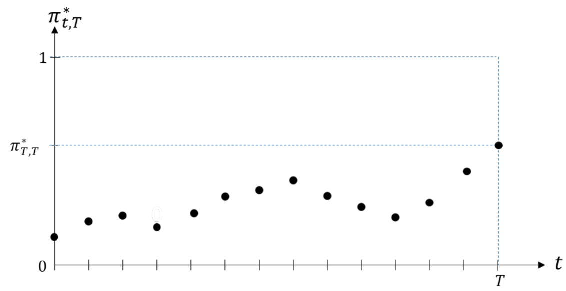
Lemma 2 (Optimal Auditing Thresholds)**.** *Optimal auditing strategies are indicator functions such that all priors above a threshold are audited with certainty and all priors below are never audited. In other words, there exists a sequence* $\left\{\pi^*_{t,T}\right\}_{t\in\{1,..,T\}} \in (0,1)^{T+1}$ *such that*

$$x^*_{t,T}(\pi)\begin{cases} = 0 & \text{if} \quad \pi < \pi^*_{t,T}, \\ \in (0,1) & \text{if} \quad \pi = \pi^*_{t,T}, \\ = 1 & \text{if} \quad \pi > \pi^*_{t,T}. \end{cases}$$

The insurer's decision can therefore be reduced to the choice of thresholds $\left\{\pi^*_{t,T}\right\}_{t\in\{1,..,T\}}$.



(a) Indicator Optimal Strategy



(b) Optimal Thresholds Sequence

In this transformed context, our goal is now to understand how the sequence of optimal auditing thresholds change across time for a fixed number of periods, and how the initial auditing thresholds $\{\pi^*_{0,T}\}_{T\in\mathbb{N}}$ evolve when the number of periods T becomes arbitrarily large, i.e,

$$\pi^*_{t,T} \gtrless \pi^*_{t+1,T},$$
$$\pi^*_{1,T} \xrightarrow[T\to\infty]{} ?.$$

3.2 The two-period models

Before studying the dynamic programming model for an arbitrary number of periods, let us look at the two-period case scrutinized in Aboutajdine & Picard (2018).

Lemma 3. *For a two-period model $T = 2$ and it is optimal to include individually unprofitable claims to audit at the first period $t = 1$. In other words*

$$\pi^*_{1,2} < \pi^*_{2,2},$$

and

$$\forall \pi \in (\pi^*_{1,2}, \pi^*_{2,2}), \bar{p}(\pi) - c < 0.$$

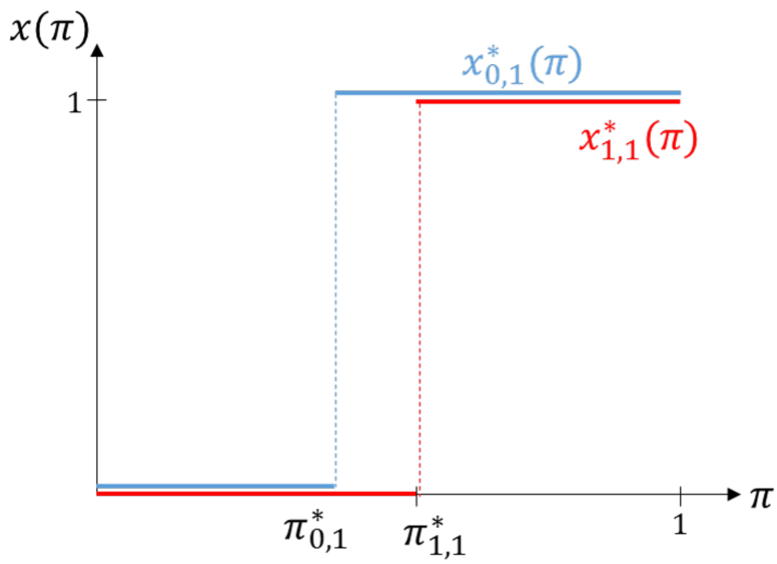


Figure 5: Two-period Auditing

This is the central result from Aboutajdine & Picard (2018), namely that early auditing, i.e., at the first period, should include more claims, even is they are individually non-profitable to audit, due to a learning effect: the early loss is compensated by larger auditing proceeds at the second period because the acquired information early on allows for a better informed audit in the future. In this case, the auditor deviates from the pure exploitation strategy to allow for some exploration, but does not change for a pure exploration strategy.

From this point, it is natural to ask two questions

- How does this expansion of the auditing set at earlier periods translate in a context with more periods?
- Can it be optimal for the auditor to switch to a pure exploration strategy at earlier periods (and if so, under what conditions)?

3.3 Optimal Inter-temporal Auditing

Proposition 1 answers our first question.

Proposition 1. *Let $(\pi^*_{1,T})_{T\in\mathbb{N}}$ be the sequence of initial period thresholds. Thresholds verify the following assertions*

i *For any $T \in \mathbb{N}$, the sequence $(\pi^*_{t,T})_{t\in[\![0,T]\!]}$ is strictly increasing.*

ii *$(\pi^*_{1,T})_{T\in\mathbb{N}}$ is strictly decreasing.*

iii $\forall i \in \{1, .., t-1\}$, $\pi^*_{t,T} = \pi^*_{t-i,T-i}$

iv *$(\pi^*_{1,T})_{T\in\mathbb{N}}$ converges towards a limit $\pi^*_{1,\infty} \in [0, \pi^+)$*

Statement (*i*) means that, for a given number of periods, the earlier the audit, the larger the auditing set. This is due to the fact that interacting longer with auditees induces the auditor to learn more about them, even when their claims are believed to be non profitable to audit at that point in time, because the better targeted audits later on are performed enough times to cover the preliminary expenses. In terms of exploration/exploitation, the auditor explores more at the beginning, hoping to identify dishonest types soon enough to exploit them later. Statement (*ii*) is a direct equivalent of statement (*i*) because of the backward induction nature of the solution. Statement (*iii*) indicated that the optimal threshold depends on the number of remaining periods. Statement (*iv*) means that there is an asymptotic optimal level of exploration, which may be pure exploration if the limit is 0.

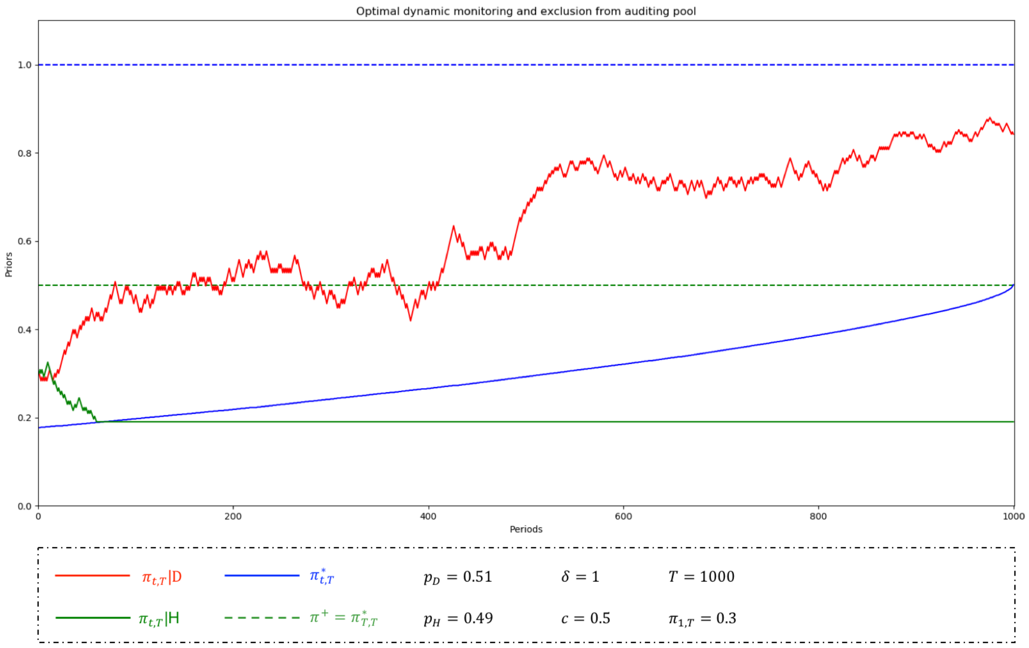


Figure 6: Optimal Threshold and Exclusion from Auditing Pool

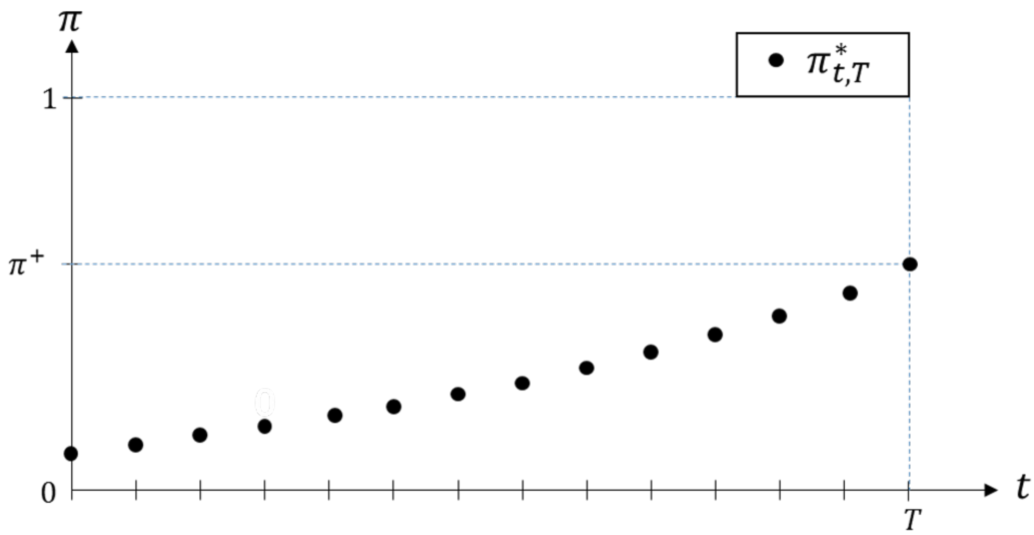


Figure 7: Increasing Thresholds

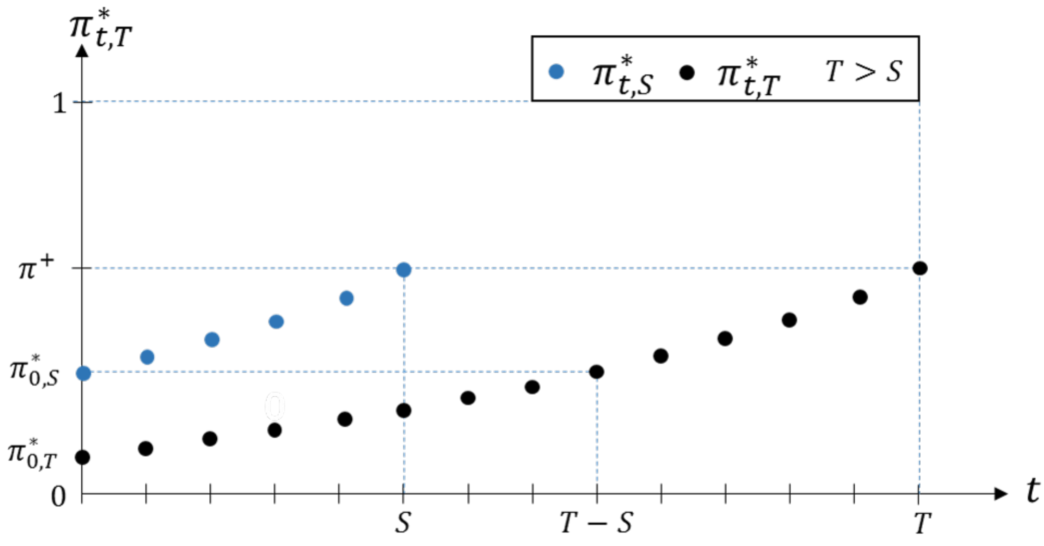


Figure 8: First Period Auditing when T increases

4 Degrees of informativeness and time preference

4.1 A Martingale Characterization of Priors

Consider the following random variables, at each period t and for a belief π_t, with three different probabilities given the state of the world $(\mathbb{P}, \mathbb{P}_D, \mathbb{P}_H)$:

$I_t \in \{0, 1\}$ a random variable equal to 1 if a claim when a claim is invalid

$\mathbb{P}(I_{t+1} = 1) = \bar{p}(\pi_t) \quad \mathbb{P}_D(I_{t+1} = 1) = p_D \quad \mathbb{P}_H(I_{t+1} = 1) = p_H$

$p_H < \bar{p}(\pi_t) < p_D$

$(\mathcal{F})_t$ the filtration generated by the observations $(I_1, .., I_t)$

Π_{t+1} a random variable corresponding to the updating process following an audit at time t

$\Pi_{t+1} = I_{t+1}A(\Pi_t) + (1 - I_{t+1})B(\Pi_t)$

Under the different states of the world, beliefs evolve in different fashions.

Proposition 2. *[Martingale Properties] Depending on the different states of the world, the prior behaves as a submartingale, a supermartingale or a martingale.*

i Under $\mathbb{P}_H$, Π_t is a supermartingale, $\mathbb{E}_H[\Pi_{t+1}|\mathcal{F}_t] \leq \Pi_t$ and $\lim \mathbb{E}_H[\Pi_{t+1}|\mathcal{F}_t] = 0$.

ii Under $\mathbb{P}$, Π_t is a martingale and $\mathbb{E}[\Pi_{t+1}|\mathcal{F}_t] = \Pi_t$.

iii Under $\mathbb{P}_D$, Π_t is a submartingale, $\mathbb{E}_D[\Pi_{t+1}|\mathcal{F}_t] \geq \Pi_t$ and $\lim \mathbb{E}_D[\Pi_{t+1}|\mathcal{F}_t] = 1$.

This martingale characterization allows us to identify the magnitude of information acquisition

Proposition 3. *[Information acquisition] Under states of the world H and D, we can decompose the stochastic process of updated beliefs thanks to the Doob-Meyer decomposition.*

i *Under* $\mathbb{P}_H$, $\Pi_t = M_t^H + P_t^H$ *where* M_t^H *is a martingale and* P_t^H *is a decreasing and predictable process.*

ii $P_t^H = -(p_D - p_H) \sum_{i=0}^{t-1} \Pi_i \Big[A(\Pi_i) - B(\Pi_i)\Big]$

iii *Under* $\mathbb{P}_D$, $\Pi_t = M_t^D + P_t^D$ *where* M_t^D *is a martingale and* P_t^D *is an increasing and predictable process.*

iv $P_t^D = (p_D - p_H) \sum_{i=0}^{t-1} \Big(1 - \Pi_i\Big) \Big[A(\Pi_i) - B(\Pi_i)\Big]$

The predictable process P_t represents the average gain in information under true states H and D.

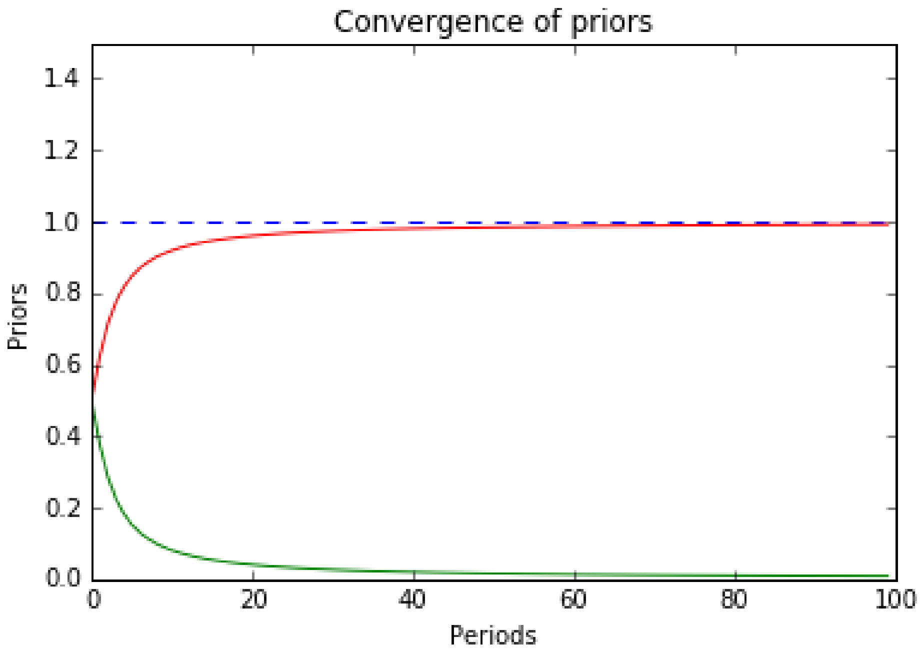


Figure 9: Conditional expectation of priors under states of the world H (green) and D (red)

4.2 Informativeness and exploration

The decomposition in Proposition 3 allows us to quantify the amount of information obtained at each stage and to understand how it affects the optimal level of exploration. From the expression of P_t, we can see that the difference $\beta = p_D - p_H$ plays an important role. This is rather intuitive in that a larger difference means more polarized behavior between the Honest and Dishonest behaviors.

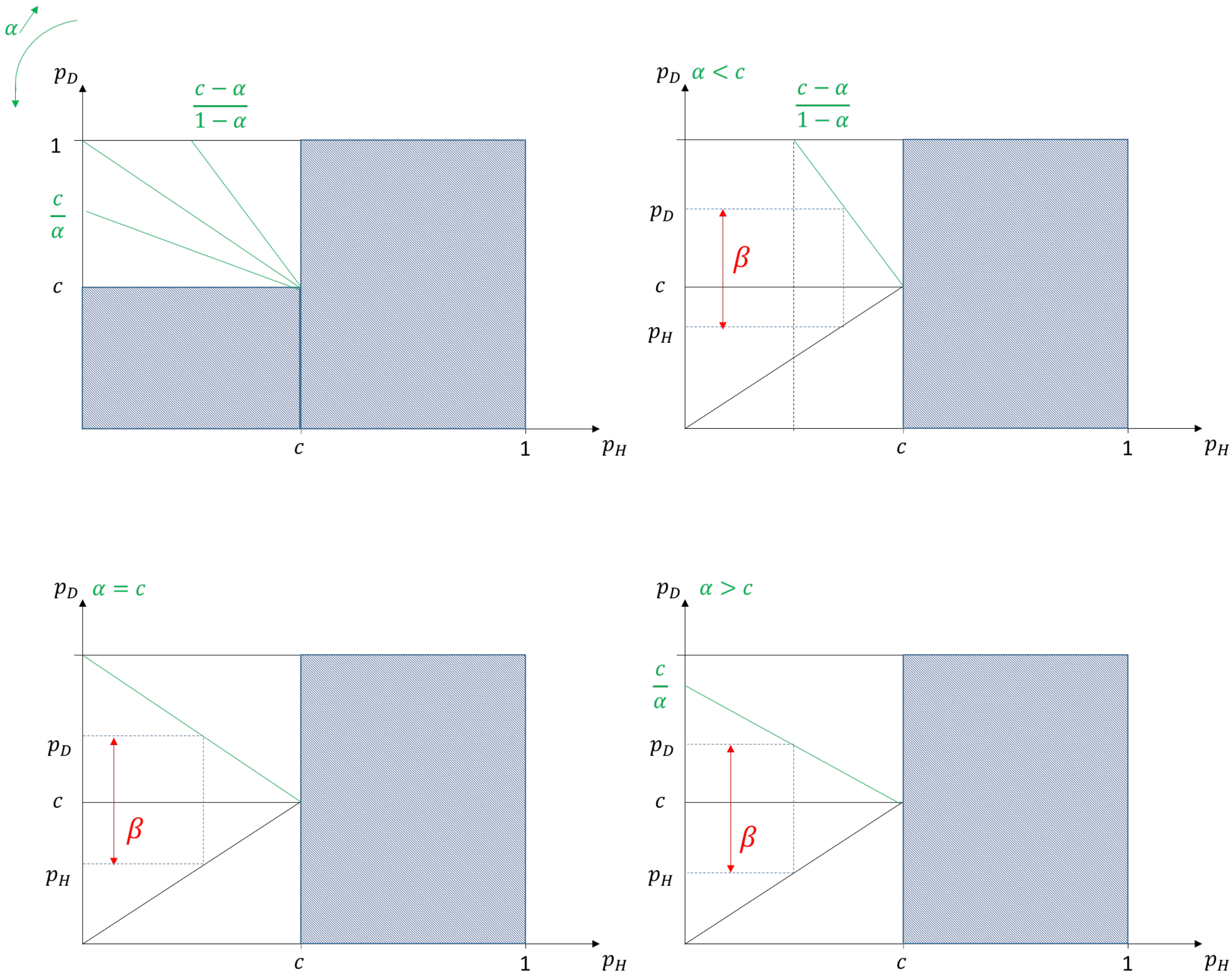


Figure 10: Variations in informativeness

Proposition 4. *At a fixed cost c and instantaneous auditing profitability α, the more informative the audit (i.e., the higher $\beta = p_D - p_H$), the higher the optimal exploration level at every period, i.e., for any t, T,*

$$\frac{\partial \pi^*_{t,T}}{\partial \beta} < 0$$

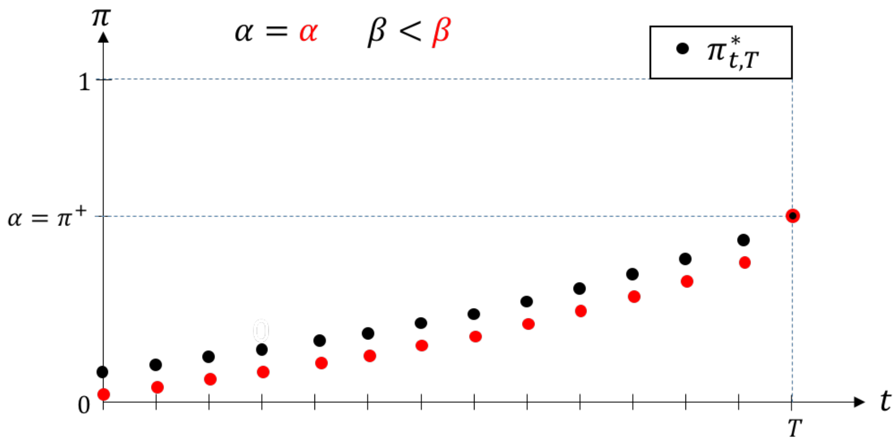


Figure 11: Exploration Levels for Different Informativeness Levels

4.3 Time preference and exploration

Proposition 5. *The function $\pi^*_{1,\infty}(\delta)$ is decreasing in δ, i.e., the more patient the insurer (i.e., the higher δ), the larger the asymptotic auditing set $(\pi^*_{1,\infty}(\delta), 1)$ at the initial period, with*

$$(i)\frac{\partial \pi^*_{1,\infty}}{\partial \delta} < 0$$
$$(ii)\lim_{\delta \to 0} \pi^*_{1,\infty}(\delta) = \pi^+ = \frac{c - p_H}{p_D - p_H},$$
$$(iii)\lim_{\delta \to 1} \pi^*_{1,\infty}(\delta) = 0.$$

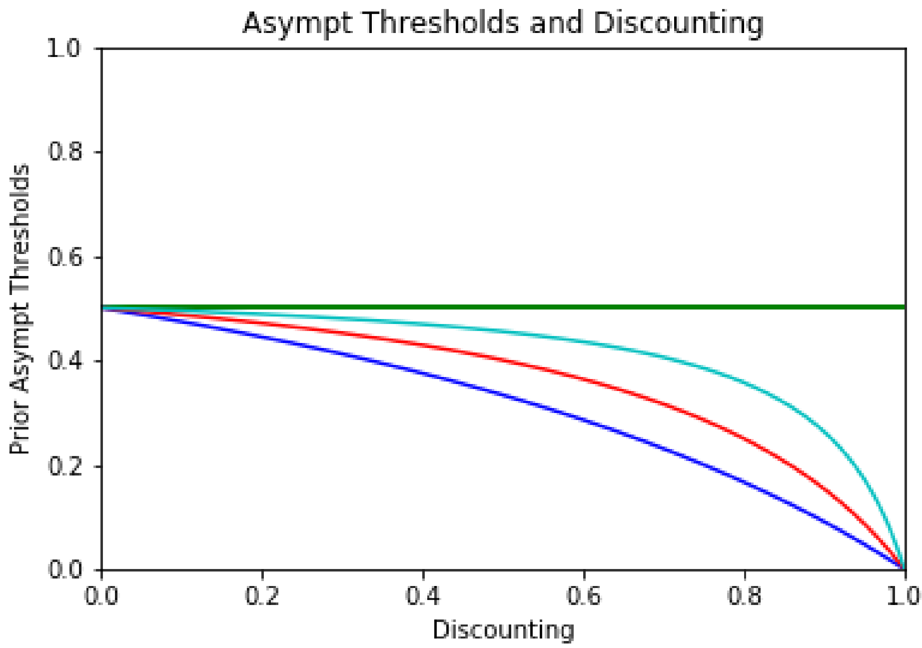


Figure 12: Asymptotic first period thresholds and time-discounting ($\pi^+ = 0.5$).

Proposition 5 states that a necessary and sufficient condition for the whole population to be audited at the initial period when T becomes arbitrarily large is for the auditor to be patient enough, i.e., δ close enough to 1. The underlying intuition is that while additional auditing allows to have refined information and better targeted audits, the lower the initial period prior, the longer the insurer has to wait before he can derive positive proceeds for the corresponding SPs.

5 Conclusion

There are many arguments upon which an insurer may base his decision to audit in the context of insurance fraud. One of them is the opportunity to refine his beliefs and target defrauders more accurately in the future.

In this article, we analyzed this decision in a finite-horizon multiperiod setting, where the Bayesian insurer could be conflicted between prioritizing immediate payoffs (exploitation) and favoring future proceeds (exploration). Interestingly, our results establish that one determining feature of the decision problem was the remaining number of periods, in which the levels of exploration are increasing. Unsurprisingly, the type-separating power of auditing and the insurer's patience were also determinants of exploration levels.

Clearly, these results are only a step in what we hope is the right direction. While they bring many answers to the questions faced by an insurer who wishes to fight fraud, the analysis will only be complete once both deterrence and information are included in the same model. This is already possible in our framework by relaxing the assumption of exogeneity for the Service Providers fraud levels. Another possible direction to explore would be the inclusion of information externalities in auditing: when the Insurer monitors an auditee, he may learn about all other agents that resemble the monitored one. Finally, an interesting practical application of our model would be to construct an exclusion process whereby a Service Provider would be excluded from an insurer's affiliation network one he becomes suspicious enough.

References

Aboutajdine, R. & Picard, P. (2018), 'Preliminary Investigations for Better Monitoring: Learning in Repeated Insurance Audits', *Risks* **6**(1), 15.

Aghion, P., Bolton, P., Harris, C. & Jullien, B. (1991), 'Optimal Learning by Experimentation', *The Review of Economic Studies* **58**(4), 621–654.

Allingham, M. G. & Sandmo, A. (1972), 'Income tax evasion: A theoretical analysis', *Journal of Public Economics* **1**(3-4), 323–338.

Artís, M., Ayuso, M. & Guillén, M. (2002), 'Detection of Automobile Insurance Fraud With Discrete Choice Models and Misclassified Claims', *Journal of Risk and Insurance* **69**(3), 325–340.

Bergemann, D. & Välimäki, J. (2008), Bandit Problems, *in* S. N. Durlauf & L. E. Blume, eds, 'The New Palgrave Dictionary of Economics: Volume 1 – 8', Palgrave Macmillan UK, London, pp. 336–340.

Brockett, P. L., Derrig, R. A., Golden, L. L., Levine, A. & Alpert, M. (2002), 'Fraud Classification Using Principal Component Analysis of RIDITs', *Journal of Risk and Insurance* **69**(3), 341–371.

Derrig, R. A. (2002), 'Insurance Fraud', *Journal of Risk and Insurance* **69**(3), 271–287.

Dionne, G., Giuliano, F. & Picard, P. (2008), 'Optimal Auditing with Scoring: Theory and Application to Insurance Fraud', *Management Science* **55**(1), 58–70.

Gale, D. & Hellwig, M. (1985), 'Incentive-Compatible Debt Contracts: The One-Period Problem', *The Review of Economic Studies* **52**(4), 647–663.

Gossner, O. (2011), 'Simple Bounds on the Value of a Reputation', *Econometrica* **79**(5), 1627–1641.

Gossner, O. & Tomala, T. (2008), 'Entropy bounds on Bayesian learning', *Journal of Mathematical Economics* **44**(1), 24–32.

Guala, F. & Mittone, L. (2005), 'Experiments in economics: External validity and the robustness of phenomena', *Journal of Economic Methodology* **12**(4), 495–515.

Kirchler, E. (2007), *The Economic Psychology of Tax Behaviour*, Cambridge University Press.

Major, J. A. & Riedinger, D. R. (2002), 'EFD: A Hybrid Knowledge/Statistical-Based System for the Detection of Fraud', *Journal of Risk and Insurance* **69**(3), 309–324.

Mittone, L. (2006), 'Dynamic behaviour in tax evasion: An experimental approach', *The Journal of Socio-Economics* **35**(5), 813–835.

Mookherjee, D. & Png, I. (1989), 'Optimal Auditing, Insurance, and Redistribution', *The Quarterly Journal of Economics* **104**(2), 399–415.

Picard, P. (2013), Economic Analysis of Insurance Fraud, *in* G. Dionne, ed., 'Handbook of Insurance', Springer New York, pp. 349–395.

Rosenberg, D., Solan, E. & Vieille, N. (2007), 'Social Learning in One-Arm Bandit Problems', *Econometrica* **75**(6), 1591–1611.

Sutton, R. S. & Barto, A. G. (1998), *Reinforcement Learning: An Introduction*, MIT Press.

Tennyson, S. & Salsas-Forn, P. (2002), 'Claims Auditing in Automobile Insurance: Fraud Detection and Deterrence Objectives', *Journal of Risk and Insurance* **69**(3), 289–308.

Townsend, R. M. (1979), 'Optimal contracts and competitive markets with costly state verification', *Journal of Economic Theory* **21**(2), 265–293.

Viaene, S., Derrig, R. A., Baesens, B. & Dedene, G. (2002), 'A Comparison of State-of-the-Art Classification Techniques for Expert Automobile Insurance Claim Fraud Detection', *Journal of Risk and Insurance* **69**(3), 373–421.

Zitzewitz, E. (2012), 'Forensic Economics', *Journal of Economic Literature* **50**(3), 731–769.

Appendix

5.1 Proofs

5.1.1 Proof of Lemma 2

Denoting $\frac{\partial \Omega_{t,T}}{\partial x_{t,T}}\Big|_{\pi} = \Omega'_{t,T}(\pi)$, notice that:

- $\Omega'_{t,T}(0) < 0$,
- $\Omega'_{t,T}(1) > 0$,
- $\frac{\partial \Omega'_{t,T}}{\partial \pi}(\pi) > 0$.

  Then $\pi^*_{t,T}$ is defined uniquely by

$$\Omega'_{t,T}(\pi^*_{t,T}) = 0.$$

5.1.2 Proof of Lemma 1 and Lemma 3

See Aboutajdine & Picard (2018)

5.1.3 Proof of Proposition 1

Proof. For a given $T \in \mathbb{N}$, denote

$$\Omega_{t,T}(\pi, \delta, x) = \Big[\bar{p}(\pi) - c\Big]x + \delta \sum_{\pi' \in U(\{\pi\})} \mathbb{P}(\pi'|\pi, x) V_{t+1,T}(\pi'),$$

then

$$\frac{\partial}{\partial x}\Omega_{t,T} = \Omega'_{t,T}(\pi, \delta) = \bar{p}(\pi) - c + \delta\Big[\bar{p}(\pi)V_{t+1,T}(A(\pi)) + (1 - \bar{p}(\pi))V_{t+1,T}(B(\pi)) - V_{t+1,T}(\pi)\Big].$$

Denoting

$$x^*_{t,T}(\pi) = \arg\max_x \left\{ \Big[\bar{p}(\pi) - c\Big]x + \delta \sum_{\pi' \in U(\{\pi\})} \mathbb{P}(\pi'|\pi, x) V_{t+1}(\pi') \right\}$$

and

$$\pi^*_{t,T} \quad s.t. \quad \Omega'_{t,T}(\pi^*_{t,T}, \delta) = 0.$$

We easily show that $\Omega'_{t,T}(\pi, \delta)$ is increasing in π and $\Omega_{t,T}(0, \delta). < 0 < \Omega_{t,T}(1, \delta)$

We now wonder whether $\pi^*_{t,T} < \pi^*_{t+1,T}$. To show the latter, it would be sufficient to show that

$$\frac{\partial}{\partial x}\Omega_{t,T}\Big|_{\pi^*_{t+1,T}} > 0.$$

Assume that $\pi^*_{t+1,T} < \pi^*_{t+2,T} < .. < \pi^*_{T,T}$.

Then, at time t

$$\frac{\partial}{\partial x}\Omega_{t,T}\Big|_{\pi^*_{t+1,T}} = \bar{p}(\pi^*_{t+1,T}) - c + \delta\Big[\bar{p}(\pi^*_{t+1,T})V_{t+1}(A(\pi^*_{t+1,T})) + (1-\bar{p}(\pi^*_{t+1,T}))V_{t+1}(B(\pi^*_{t+1,T})) - V_{t+1}(\pi^*_{t+1,T})\Big]$$

By definition of $\pi^*_{t+1,T}$

$$\begin{aligned}\frac{\partial}{\partial x}\Omega_{t+1}\Big|_{\pi^*_{t+1,T}} &= \bar{p}(\pi^*_{t+1,T}) - c + \delta\Big[\bar{p}(\pi^*_{t+1,T})V_{t+2}(A(\pi^*_{t+1,T})) + (1-\bar{p}(\pi^*_{t+1,T}))V_{t+2}(B(\pi^*_{t+1,T})) - V_{t+2}(\pi^*_{t+1,T})\Big] \\ &= 0\end{aligned}$$

Then, subtracting the latter from the previous

$$\begin{aligned}\frac{\partial}{\partial x}\Omega_{t,T}\Big|_{\pi^*_{t+1,T}} = \delta\Big[&\bar{p}(\pi^*_{t+1,T})(V_{t+1} - V_{t+2})(A(\pi^*_{t+1,T})) + (1-\bar{p}(\pi^*_{t+1,T}))(V_{t+1} - V_{t+2})(B(\pi^*_{t+1,T})) \\ &- (V_{t+1} - V_{t+2})(\pi^*_{t+1,T})\Big].\end{aligned}$$

In addition, for any π

$$V_{t+1}(\pi) > V_{t+2}(\pi) \tag{2}$$

since, at time $t+1$, playing the sequence $(x^*_{t+1}, x^*_{t+2}, .., x^*_{T-1}, x^*_T)$ gives $V_{t+1}(\pi)$ and dominates any other sequence, in particular $(x^*_{t+2}, x^*_{t+3}, .., x^*_T, 0)$ that gives $V_{t+2}(\pi) \geq 0$.

Finally, since $B(\pi^*_{t+1,T}) < \pi^*_{t+1,T} < \pi^*_{t+2}$

$$V_{t+1}(\pi^*_{t+1,T}) = V_{t+2}(\pi^*_{t+1,T}) = V_{t+1}(B(\pi^*_{t+1,T})) = V_{t+2}(B(\pi^*_{t+1,T})) = 0$$

In the end, from equation (2)

$$\frac{\partial}{\partial x}\Omega_{t,T}\Big|_{\pi^*_{t+1,T}} = \delta\bar{p}(\pi^*_{t+1,T})(V_{t+1} - V_{t+2})(A(\pi^*_{t+1,T})) > 0$$

□

We prove statement (ii) using statement (i) and the backward induction property (for any $i \in \{0, 1, .., t\}$, $\pi^*_{t,T} = \pi^*_{t-i,T-i}$),

$$\pi^*_{0,T+1} < \pi^*_{1,T+1} = \pi^*_{0,T}.$$

Statement (iii) is straightforward since the sequence $(\pi^*_{0,T})_{T\in\mathbb{N}}$ is bounded below by 0, which implies that it converges to a limit $\pi^*_{0,\infty} \in [0, \pi^+)$.

5.1.4 Proof of Proposition 2

Proof. At period $t+1$

$$\begin{aligned}
\mathbb{E}[\Pi_{t+1}|\mathcal{F}_t] &= \mathbb{E}[I_{t+1}A(\Pi_t) + (1 - I_{t+1})B(\Pi_t)|\mathcal{F}_t] \\
&= \mathbb{E}[I_{t+1}|\mathcal{F}_t]A(\Pi_t) + \mathbb{E}[(1 - I_{t+1})|\mathcal{F}_t]B(\Pi_t) \\
&= \mathbb{E}[\mathbb{1}_{\{I_{t+1}=1\}}|\mathcal{F}_t]A(\Pi_t) + \mathbb{E}[\mathbb{1}_{\{I_{t+1}=0\}}|\mathcal{F}_t]B(\Pi_t) \\
&= \mathbb{P}(I_{t+1} = 1|\mathcal{F}_t)A(\Pi_t) + \mathbb{P}(I_{t+1} = 0|\mathcal{F}_t)B(\Pi_t) \\
&= \bar{p}(\Pi_t)A(\Pi_t) + (1 - \bar{p}(\Pi_t))B(\Pi_t) \\
&= \Pi_t
\end{aligned}$$

$$\begin{aligned}
\mathbb{E}_D[\Pi_{t+1}|\mathcal{F}_t] &= \mathbb{E}_D[I_{t+1}A(\Pi_t) + (1 - I_{t+1})B(\Pi_t)|\mathcal{F}_t] \\
&= \mathbb{E}_D[I_{t+1}|\mathcal{F}_t]A(\Pi_t) + \mathbb{E}_D[(1 - I_{t+1})|\mathcal{F}_t]B(\Pi_t) \\
&= \mathbb{E}_D[\mathbb{1}_{\{I_{t+1}=1\}}|\mathcal{F}_t]A(\Pi_t) + \mathbb{E}_D[\mathbb{1}_{\{I_{t+1}=0\}}|\mathcal{F}_t]B(\Pi_t) \\
&= \mathbb{P}_D(I_{t+1} = 1|\mathcal{F}_t)A(\Pi_t) + \mathbb{P}_D(I_{t+1} = 0|\mathcal{F}_t)B(\Pi_t) \\
&= p_D A(\Pi_t) + (1 - p_D)B(\Pi_t) \\
&= \Pi_t + \Big(p_D - \bar{p}(\Pi_t)\Big)\Big[A(\Pi_t) - B(\Pi_t)\Big] \\
&= \Pi_t + (1 - \Pi_t)\Big(p_D - p_H\Big)\Big[A(\Pi_t) - B(\Pi_t)\Big] \\
&\geq \Pi_t
\end{aligned}$$

$$\begin{aligned}
\mathbb{E}_H[\Pi_{t+1}|\mathcal{F}_t] &= \mathbb{E}_H[I_{t+1}A(\Pi_t) + (1 - I_{t+1})B(\Pi_t)|\mathcal{F}_t] \\
&= \mathbb{E}_H[I_{t+1}|\mathcal{F}_t]A(\Pi_t) + \mathbb{E}_H[(1 - I_{t+1})|\mathcal{F}_t]B(\Pi_t) \\
&= \mathbb{E}_H[\mathbb{1}_{\{I_{t+1}=1\}}|\mathcal{F}_t]A(\Pi_t) + \mathbb{E}_H[\mathbb{1}_{\{I_{t+1}=0\}}|\mathcal{F}_t]B(\Pi_t) \\
&= \mathbb{P}_H(I_{t+1} = 1|\mathcal{F}_t)A(\Pi_t) + \mathbb{P}_H(I_{t+1} = 0|\mathcal{F}_t)B(\Pi_t) \\
&= p_H A(\Pi_t) + (1 - p_H)B(\Pi_t) \\
&= \Pi_t + \Big(p_H - \bar{p}(\Pi_t)\Big)\Big[A(\Pi_t) - B(\Pi_t)\Big] \\
&= \Pi_t - \Pi_t\Big(p_D - p_H\Big)\Big[A(\Pi_t) - B(\Pi_t)\Big] \\
&\leq \Pi_t
\end{aligned}$$

□

5.1.5 Proof of Proposition 3

Proof. The updated prior under a given state of the world $\{H\}$ or $\{D\}$ can be written as

$$\begin{aligned}
\Pi_{t+1} &= I_{t+1}A(\Pi_t) + (1 - I_{t+1})B(\Pi_t) \\
&= \Pi_t + \Big(I_{t+1} - \bar{p}(\Pi_t)\Big)\Big[A(\Pi_t) - B(\Pi_t)\Big]
\end{aligned}$$

In particular

$$\begin{aligned}\mathbb{E}_D[\Pi_{t+1}|\mathcal{F}_t] &= \Pi_t + \Big(p_D - \bar{p}(\Pi_t)\Big)\Big[A(\Pi_t) - B(\Pi_t)\Big] \\ &= \Pi_t + \Big(1-\Pi_t\Big)\Big(p_D - p_H\Big)\Big[A(\Pi_t) - B(\Pi_t)\Big]\end{aligned}$$

and

$$\begin{aligned}\mathbb{E}_H[\Pi_{t+1}|\mathcal{F}_t] &= \Pi_t + \Big(p_H - \bar{p}(\Pi_t)\Big)\Big[A(\Pi_t) - B(\Pi_t)\Big] \\ &= \Pi_t - \Pi_t\Big(p_D - p_H\Big)\Big[A(\Pi_t) - B(\Pi_t)\Big]\end{aligned}$$

Finally, the decomposition, itself is given by (under state of the world $\{D\}$)

$$\begin{aligned}&\Pi_t = M_t + P_t \\ &M_t \text{ is a martingale and } P_t \text{ is predictable and increasing} \\ &P_t = (p_D - p_H)\sum_{i=0}^{t-1}\Big(1-\Pi_i\Big)\Big[A(\Pi_i) - B(\Pi_i)\Big] \\ &M_0 = \Pi_0 \\ &P_0 = 0\end{aligned}$$

We can verify that M_t is indeed a Martingale by induction (it is verified at $t = 1$)

$$\begin{aligned}\mathbb{E}_D[M_{t+1}|\mathcal{F}_t] &= \mathbb{E}_D[\Pi_{t+1} - P_{t+1}|\mathcal{F}_t] \\ &= \mathbb{E}_D[\Pi_{t+1}|\mathcal{F}_t] - P_{t+1} \\ &= \mathbb{E}_D[I_{t+1}A(\Pi_t) + (1-I_{t+1})B(\Pi_t)|\mathcal{F}_t] - P_{t+1} \\ &= \Pi_t + \underbrace{\Big(1-\Pi_t\Big)\Big(p_D - p_H\Big)\Big[A(\Pi_t) - B(\Pi_t)\Big] - P_{t+1}}_{=-P_t} \\ &= \Pi_t - P_t \\ &= M_t\end{aligned}$$

□

5.1.6 Proof of Proposition 5

The limit $\Omega_{0,\infty}(\pi, x, \delta) = \lim_{T\to\infty}\Omega_{0,T}(\pi, x, \delta)$ is such that

$$\begin{aligned}\Omega_{0,\infty}(\pi, 1, 0) < 0, \\ \Omega_{0,\infty}(\pi, 1, 1) > 0.\end{aligned}$$

The first inequality is straightforward since, for any $\pi < \pi^*_{0,0}$

$$\Omega_{0,T}(\pi, 1, 0) = \bar{p}(\pi) - c < 0$$

The second inequality comes from the fact that with no discount, the proceeds of auditing will be strictly positive at each period beyond a threshold T_s under the auditing strategy with systematic auditing. Therefore, $\Omega_{0,\infty}(\pi, 1, 1)$ is higher since proceeds of auditing after period 0 are maximized.

Then, by continuity in δ of $\Omega_{0,\infty}(\pi, x, \delta)$, there exists a threshold $\delta_s \in (0, 1)$ such that

$$\Omega_{0,\infty}(\pi, 1, \delta_s) > 0,$$

which would mean that the considered π is included in the auditing set. Since this is proved for any $\pi \in (0, 1)$, it implies $\pi_{0,\infty} = 0$ for $\delta > \delta_S$.

For any π, the expected gains from auditing between periods 0 and T admit as an upper-bound

$$\begin{aligned}
\underline{\Omega}_{0,T}(\pi, 1, \delta) &= \bar{p}(\pi) - c + \sum_{i=1}^{T} \delta^i (p_D - c), \\
&= \bar{p}(\pi) - c + (p_D - c)\delta \frac{1 - \delta^T}{1 - \delta} \\
\underline{\Omega}_{0,\infty}(\pi, 1, \delta) &= \bar{p}(\pi) - c + (1 - c)\frac{\delta}{1 - \delta}
\end{aligned}$$

This is indeed the maximal possible outcome of auditing systematically at each period when the auditor discovers as soon as the second period that the SP under scrutiny is a dishonest one. This function is increasing in π and δ. It defines $\underline{\pi}_\infty(\delta)$ and $\underline{\delta}$ as

$$-c + (1 - c)\frac{\underline{\delta}}{1 - \underline{\delta}} = 0 \Leftrightarrow \underline{\delta} = c,$$

and, denoting $\pi_-(\delta)$ such that

$$\bar{p}(\pi_-(\delta)) - c + (1 - c)\frac{\delta}{1 - \delta} = 0,$$

$$\underline{\pi}_\infty(\delta) = \max\Big(0, \pi_-(\delta)\Big).$$