Essays in Microeconomic Theory

Inauguraldissertation

zur Erlangung des Grades eines Doktors
der Wirtschafts- und Gesellschaftswissenschaften

durch

die Rechts- und Staatswissenschaftliche Fakultät der
Rheinischen Friedrich-Wilhelms-Universität Bonn

vorgelegt von

Jan Knoepfle

aus Donaueschingen

Bonn, 2020

Dekan:	Prof. Dr. Jürgen von Hagen
Erstreferent:	Prof. Sven Rady, PhD
Zweitreferent:	Prof. Dr. Stephan Lauermann

Tag der mündlichen Prüfung: 29. Oktober 2020

Ersatzprüfer:	Prof. Francesc Dilmé, PhD

A Laura

Acknowledgements

In preparing this thesis, I was lucky to enjoy the company of and support by numerous people. First, I want to thank my supervisor Sven Rady. As a teacher, he has been a role model since the first day of the PhD programme. As a supervisor, his continual support and advice, as well as the freedom to pursue my research interests without constraints, have led to this thesis. I am grateful to my supervisor Stephan Lauermann for the guidance and feedback that I received individually and for the time and effort invested into fostering an invaluable group dynamic among all of his students. I feel privileged to form part of this group. Parts of this thesis were written during my research visit at Yale. I am indebted to my local supervisor Marina Halac for the exceptional feedback, support, and encouragement well beyond the visit. Our countless meetings considerably shaped Chapters 1 and 3 of this thesis and provided a source of fun and motivation.

I thank the theory group at the University of Bonn; in particular, Francesc Dilmé, who was always available with useful suggestions, and Benny Moldovanu, whose research course and feedback were essential for Chapter 2 of this thesis. I thank the Department of Economics at Yale for their hospitality, and the theory faculty for stimulating conversations. My visit would not have been the same without the company of Ian Ball, Tan Gan, Patrick Lahr, Weicheng Min, and Allen Vong.

Deniz Kattwinkel and Peter A. Wagner are both great coauthors; I would like to thank them for showing me that joint work is not only more productive but also a lot of fun. Andre Speit and Lucas ter Steege went from exceptional office mates to close friends and I thank them, Paul Voß, and many other colleagues at BGSE for making graduate life a fun experience. I thank Niklas Freier, Anna Schäfer, and Michelle Trimborn for everyday life support and much-needed distraction.

My undergraduate teachers deserve special mention. Thomas Blum provided invaluable stimulus and advice to embark on an academic path. Ursula Höpping sparked and fostered my interest in economics. Without their early guidance and encouragement, I would never have envisioned myself pursuing a PhD.

Finally, I am eternally grateful to Laura for the love and encouragement during the ups and downs. This thesis would not have been possible without and is dedicated to you. I am indebted to both our families for their unconditional support.

Contents

Introduction

This thesis comprises four self-contained essays in economic theory studying the optimal use of private information in strategic interactions. In all the situations considered, non-monetary as well as dynamic incentives play an important role. Chapter 1 contributes to the theory of information design by analysing how information is optimally released to attract attention over time. The questions studied in Chapters 2-4 belong to the theory of mechanism design and contract theory. These chapters aim to understand how information is used to support incentive provision and when it is most effective to acquire costly information.

Chapter 1, 'Dynamic Competition for Attention', models competing information sources that release information over time to maximise viewership by a time-constrained consumer. When the source is a monopolist, the chapter shows that information arrives in jumps and the analysis uncovers a novel driver for the optimality of jumps: the lack of intertemporal commitment combined with the non-concavity in the value of information. For multiple senders, an equilibrium is characterised in which all information is transmitted in minimal time. The equilibrium reveals kinship between this market for information and sequential oligopolistic price competition for standard consumption goods. Information as a good has two noteworthy features also present in oligopolistic competition: capacity constraints and consumption externalities. Capacity constraints result from the sources' fixed endowments with information and – as in standard oligopolies – prevent profits from being competed away. Consumption externalities arise because each observation by the consumer affects her demand for further information as well as the expectation about unknown content. In equilibrium, each source's profit is determined by the marginal contribution of her information to the consumer's knowledge.

In Chapter 2, 'Costless Information and Costly Verification: A Case for Transparency', joint work with Deniz Kattwinkel, we study the role of correlated information in an allocation problem without money where the principal's optimal choice depends on an agent's private information. For the principal, a private correlated signal is valuable because it induces belief-heterogeneity; the private information of the agent comes with different beliefs about the signal of the principal. Making allocation decisions contingent on this signal creates heterogeneity in the expected value across agent types. With money, this contingency enables the principal to elicit the agent's private type at no cost. We show that this is not the case when monetary transfers are replaced with costly verification. In this case, transparent procedures are optimal. The principal's payoff is the same as if her signal was revealed to the agent prior to their interaction. To explain this contrast, we dissect two channels through which the principal might benefit from the secrecy of her signal: the reuse of excess utility for incentives across different signals and the increase of perceived verification risk for a specific type due to subjective beliefs. We prove pointwise monotonicity of optimal allocations in the type and the signal. Without money, these properties respectively render the first and second channel futile. We conclude that secrecy cannot be justified solely on grounds of efficiency.

The remaining two chapters study costly inspections in dynamic environments. Chapter 3, 'Inspecting Experimentation', analyses how a venture capitalist optimally combines bonus payments and inspections to incentivise a cash-constrained entrepreneur to experiment on a project of unknown quality. Bonuses and inspections have to deter the entrepreneur from diverting the funds provided by the investor. The long-term nature of this financing relationship leads to dynamic rents for the agent, which grow exponentially in the length of the contract. Inspections serve to decrease these rents by reducing the maximal amount of funds that can be diverted before being detected. When the inspection technology is precise enough, an additional benefit arises as the principal can reuse rents across time: the threat of termination serves to incentivise the entrepreneur in early periods without offering additional payments in these periods. In this case, the optimal timing of inspections is predictable rather than random. This is due to the effective discounting in experimentation relationships that end with the arrival of the first success. The discounting-induced preferences over time-lotteries imply that, from the perspective of a diverting agent, the threat of inspections is most powerful when its timing is predictable.

Chapter 4, 'Dynamic Incentives with Costly Inspections', joint work with Peter A. Wagner, studies how a compliance manager uses inspections and fines to achieve maximal compliance by an agent. In a long-term interaction, the observability of inspections facilitates a fully compliant equilibrium even when the principal has no commitment power. Without observability, incentivising the principal to inspect necessarily requires instances of non-compliance. We show further that, without commitment, the principal cannot benefit from random inspections to implement full compliance because the continuation equilibrium cannot be used to punish skipping an inspection if it was not foreseen to happen with certainty. In comparison to Chapter 3, the relation between the agent's risk-preferences on and off the equilibrium path is reversed because fulfilling his task – compliance – does not increase the probability of termination. Thus, with commitment power, random inspections weakly dominate predictable rules. We discuss possible arrangements outside our model which provide the compliance manager with the incentives or commitment power necessary to benefit from randomisation. We conclude that the separation of inspection planning and execution as observed in banking supervision is most promising.

1

Dynamic Competition for Attention

1.1 Introduction

Information providers compete for attention. Most online content, such as news, professional product reviews, or weather forecasts, is offered free of charge, and the websites exploit the attracted attention to create revenue, primarily through advertisements. Without monetary prices, information providers compete through two key factors that determine their profits. First, they have to decide how much and what type of information to acquire. Several papers studying this question highlight the importance of attention as the currency in media markets.[1] Second, providers have to decide how to reveal their information over time. This aspect – how to optimally disseminate information when competing for attention – is the focus of this paper.

Attention is collected from a decision maker who is interested in the information held by the providers. As attention requires time and effort, the decision maker decides sequentially which providers to visit and when to stop, depending on the information previously observed. Recent work has studied the design of optimal dynamic information policies from the perspective of the decision maker.[2] Yet, in many situations, the power to design how information is revealed over time lies with providers.

How much information can be transmitted from the providers to the decision maker and what type of information processes arise when providers design offers to attract attention?

To answer these questions, I build a dynamic model in which information providers – the senders – compete for the attention of a decision maker – the receiver. The receiver has to take an action and wants to learn about an unknown state to maximise his utility. Senders are interested in maximising the number of visits and

[1] See Galperti and Trevino (2018), Perego and Yuksel (2018), and Pant and Trombetta (2019).

[2] Most notably, Zhong (2019) characterises the optimal process designed by the decision maker with full flexibility. See discussion below.

do not care about the receiver's action. At the beginning of the game, each sender is endowed with imperfect information over the state through a signal. Subsequently, there are multiple rounds of communication in which senders compete for a visit by the receiver. At the beginning of each round, senders offer experiments over their signals, that is, each sender commits to a distribution over messages, conditional on the realisation of her signal. Senders cannot commit across rounds. The receiver observes all offers. He either pays an attention cost to visit one sender and continue to the next round, or he stops learning and takes the optimal action with the current level of information. The model captures broad information and preference specifications with the condition that attention can be split finely enough and each sender's signal is informative enough so that it is worth at least one unit of attention, independent of the information previously delivered by her competitors.

The main result characterises an equilibrium in which all information is transmitted from the senders to the receiver. Each sender attracts attention proportional to the expected residual value of her information. This is a lower bound on attention for each sender. Therefore, this equilibrium is receiver-preferred, information is transmitted in minimal time. Offers made in equilibrium are of a simple class: each sender posts a probability with which the experiment reveals her initial signal fully. With the remaining probability, the experiment delivers no information. I refer to this class as *All-or-Nothing* (AoN) offers.

The market for information considered in this paper features intertemporal externalities. The information observed at any sender changes the receiver's valuation for future information as well as his probability assessment of other senders' signals. Furthermore, the design of an offer and its cost (in terms of attention) are closely intertwined. With externalities and in the absence of prices or general contracts, the existence of an efficient equilibrium is not a foregone conclusion.[3] To construct the equilibrium mentioned above, this paper introduces a substitute condition on the senders' signals that requires that any sender's information is more valuable when her competitors have revealed less.

To gain intuition on the equilibrium and the class of AoN offers, consider the case of a single sender. What is the maximal amount of attention a monopolist can extract from the receiver? The receiver is willing to pay a total attention cost equal to the difference in expected utility from taking the action with or without the sender's information. Due to the lack of intertemporal commitment, the sender cannot simply require the receiver to visit her for a fixed number of rounds and then reveal all her information at the last visit. In general, a non-committed monopolist cannot give out more information than necessary to make the receiver indifferent

[3] As shown by the example in Section 1.6, externalities may impede information transmission entirely.

between spending another round of attention and taking the action at the current information. A simple way to keep the receiver indifferent is to make an AoN offer as introduced above. The sender chooses the AoN probability that all information is revealed as low as possible so that the receiver accepts. AoN offers imply that no information is revealed until a geometrically distributed arrival time, at which time all information is revealed.

When there are multiple senders, they design experiments facing Bertrand competition in every round. Each sender offers an experiment that makes her indifferent between being accepted and the lower bound of attention she can attract if she is not visited. This lower bound consists of waiting until all competitors have revealed their information and, subsequently, playing the strategy of the monopolist. At this point, the receiver's information includes all signals of her competitors, and the value of the lower bound depends on the realisations of these signals. In equilibrium, every sender offers the AoN probability such that the expected attention is equal to the current expectation of her lower bound. This expectation and the senders' offers change over time. Once only one sender is left, the receiver is indifferent between stopping and accepting this last sender's offer. As signals are substitutes, the receiver strictly prefers to accept an offer when there are still multiple senders whose information he has not yet observed. Given that senders require attention proportional to the residual value of their signal, concentrating a fixed amount of information on fewer senders hurts the receiver. While all information is still transmitted, the total required attention increases.

I provide examples of information and preference specifications that are captured by the model together with applications that the literature has studied with these specifications. Among these is the application of the Gaussian-information, quadratic-loss specification to examine competition in news markets. For this setup, I extend the game to consider optimal information acquisition by two competing news outlets that face a tradeoff between checking further sources more carefully and breaking the news as early as possible. This *investigation race* always leads to specialisation into a less informed outlet that offers a more superficial report early and a more informed outlet that investigates as long as possible to deliver high precision. If news outlets have different efficiency levels ex ante, i.e. different rates at which their precision increases over time, the more efficient newspaper is the one that investigates longer, thereby exacerbating its initial advantage. Perhaps surprisingly, increasing the precision of initially available public information may decrease the final precision at which the action is taken. The adverse effect on the incentives to investigate can outweigh the direct increase in precision. If the government considers increasing information on an issue through a campaign and ignores the incentives of other providers informing on the same issue, such campaigns may have the opposite effect and decrease information to the public.

After discussing the related literature below, Section 1.2 presents the model. Section 1.3 sets the stage for the analysis, examining the value of information and introducing useful notation. The equilibrium characterisation in Section 1.4 starts with the monopoly benchmark before deriving the results for multiple senders. The news application is considered in Section 1.5. Section 1.6 discusses modelling choices and the relation to Zhong (2019) in more detail. Concluding remarks are presented in Section 1.7. Proofs not included in the main text can be found in the Appendix.

Related Literature. This paper contributes to the literature on optimal dynamic information acquisition by a decision maker, firstly, by endogenising the information processes chosen by senders, and secondly, by considering attention maximisation. The tractable dynamic model with multiple senders who are partially informed presents a technical contribution to the dynamic information design literature. With the application to news markets, the analysis sheds light on the tradeoff between publishing news earlier or gathering more precise information. The relation to these three strands of literature, among others, is discussed in detail below.

Optimal **dynamic information acquisition** by a decision maker has been introduced to the economics literature by Wald (1947), where the decision maker decides when to stop observing an exogenous information process and take an action. Several papers enrich the decision maker's problem by allowing him to adjust the information intensity or to choose among exogenous processes, see Moscarini and Smith (2001), Mayskaya (2017), Che and Mierendorff (2019), Liang and Mu (2020), Liang et al. (2019), and others.[4] The decision maker in my paper faces a related acquisition problem but chooses among experiments that are offered endogenously by the senders. More recently, Zhong (2019) characterises the optimal information process designed by a decision maker with full flexibility facing a precision cost. He shows that the optimal policy consists of a Poisson process that leads to immediate action after arrival.[5] The current paper contributes to the literature on optimal dynamic design by considering a related question from the opposite perspective. Senders choose flexibly how to provide their information over time to maximise the attention they attract from the receiver. The offer strategies presented in the current paper[6] imply a geometrically distributed arrival of all information from one sender. This is akin to a Poisson process in continuous time with fully revealing news and where the absence of arrival allows no inference (there is no belief drift). Section

[4]See also Morris and Strack (2017) and Fudenberg et al. (2018) for recent developments on the Wald problem in different directions.

[5]This gives a theoretical justification for the common use of Poisson processes to model information in dynamic environments, partly due to its tractability laid out in Keller et al. (2005).

[6]Which attain the unique equilibrium payoff in the monopoly case and the receiver-preferred equilibrium with competition.

1.6.2 discusses the connection between Zhong (2019) and the single sender case in the current paper. I identify the lack of intertemporal commitment as an additional motive for Poisson processes.

My paper is related to the literature on **Bayesian persuasion**, based on Kamenica and Gentzkow (2011), in which a sender designs information to influence a receiver's behaviour. Senders in my model maximise attention and have no persuasion motive, that is, the action eventually taken by the receiver does not affect the senders' utilities. Most contributions model information design using belief-based techniques. I use an experiment-based approach. Section 1.6.1 discusses the benefits of doing so for a setting with multiple senders who design how to reveal partial information.

Dynamic information design has been studied in Au (2015), Che and Hörner (2017), Ely (2017), Renault et al. (2017), Smolin (2017), Board and Lu (2018), Ball (2019), Che et al. (2020), Ely and Szydlowski (2020), Guo and Shmaya (2019), Orlov et al. (2020), and others. The main contrasts to the current paper are the persuasion motive mentioned above and the focus on single-sender[7] environments. Ely and Szydlowski (2020) study the problem of a sender with intertemporal commitment who wants to persuade the receiver to execute an option as early as possible or as late as possible. The latter case may be interpreted as paying attention for as long as possible before stopping, which relates to the single sender case in the current paper. The optimal information processes in both papers share similar features: the receiver's belief is kept constant, and he is indifferent between stopping and continuing until the information is fully revealed and he stops. As in Au (2015) and Che et al. (2020), senders in my paper commit to the information offered within a period but cannot commit across periods. The senders' focus on the receiver's costly attention connects the current paper and Che et al. (2020), who examine optimal dynamic persuasion when the receiver has to pay an attention cost.

Section 1.5 considers a concrete specification with Gaussian information and quadratic loss for the receiver. This tractable setting is widely used in the literature on **media competition**. This application to news markets is related to Mullainathan and Shleifer (2005), Besley and Prat (2006), Gentzkow and Shapiro (2006), Galperti and Trevino (2018), Perego and Yuksel (2018), and Pant and Trombetta (2019). These papers highlight the importance of capturing an audience or maximising attention for media companies. They study aspects from information acquisition to optimal provision. Mullainathan and Shleifer (2005), Gentzkow and Shapiro (2006), and Pant and Trombetta (2019) consider optimal provision, which is the

[7] With the exception of Board and Lu (2018) who consider multiple sellers which are randomly matched with potential buyers in a search market and want to induce buyers to buy from them rather than continuing to search. For a cheap talk model that features multiple senders in a dynamic environment, see Margaria and Smolin (2018).

main focus of the current paper. Yet, within the more tractable Gaussian setting, I extend the analysis and consider information acquisition where precision levels arise from a market entry game. News outlets face a tradeoff between investigating longer or breaking a story before the competition. Galperti and Trevino (2018) and Perego and Yuksel (2018) study optimal static acquisition decisions determined by competition for attention. Senders in Galperti and Trevino (2018) choose the accuracy and the clarity of their news pieces at a cost, while receivers have a coordination motive. Senders in Perego and Yuksel (2018) choose to report on more polarised issues due to competition. The market entry game in the current paper focuses on accuracy levels only, and I find that the investigation race to publish first leads to polarisation in accuracy, where the less productive newspaper reports early and the more productive newspaper deepens her advantage by investigating as long as possible.

One interpretation of the setting in the current paper is to view senders as selling information to a receiver requiring a price in units of attention time. For a survey on **markets for information**, see Bergemann and Bonatti (2019). Bergemann et al. (2018) study the optimal design and pricing of a menu of experiments to screen the receiver according to his willingness to pay. In my paper, the receiver's willingness to pay is known. The monopolist can require attention proportional to the price charged by the monopolist in Bergemann et al. (2018), who would only offer the fully revealing experiment if he knew the receiver's willingness to pay. The dynamic information market studied in the current paper shares two important features with **dynamic price competition** for standard goods. First, the product 'information' has important externalities. Bergemann and Välimäki (2006) study repeated Bertrand price competition with externalities – the surplus of each purchase depends on the history of previous purchases. For the special case without inter-group externalities – where the surplus generated by a trade with seller i depends only on the number previous trades with i – they show that a marginal contribution equilibrium exists and leads to efficiency. Information generally features inter-group externalities. Nevertheless, the receiver-preferred equilibrium presented in the current paper is also constructed by considering each sender's expected marginal contribution. The second noteworthy feature of this market is the presence of capacity constraints: Each sender is initially endowed with information. This fixed endowment is a capacity constraint, relating the current paper to Dudey (1992), Martínez-de-Albéniz and Talluri (2011), and Anton et al. (2014). As information is assumed to be always worth one visit in the current paper, the capacity constraint is binding. Paralleling results in the above papers, this implies that each sender can extract positive surplus despite the competition.

Gossner et al. (2019) study attention with a different focus. They show that drawing attention to one of several considered options unequivocally increases the likelihood of this option being chosen by an agent who uses a fixed threshold rule. This can be seen as an additional rationale to compete for attention.

1.2 Model

1.2.1 Environment

One receiver and a finite number of senders $i \in \{1, \cdots, I\}$ interact in discrete rounds of communication, $k = 0, 1, \ldots$. There is a state of the world ω from the Polish[8] space Ω that remains constant and is not observed by any player. However, each sender i is endowed with partial information over the state, represented by a signal x_i from the Polish space X_i. Let $X = \times_i X_i$. The state and signals are jointly distributed according to the commonly known prior $\tilde{\mu} \in \Delta(\Omega \times X)$.

The receiver has to take an irrevocable action a from the closed space A to maximise his utility $u(a, \omega)$, where $u : A \times \Omega \to \mathbb{R}$ is continuous and bounded. For this, the receiver relies on information from the senders. In each round, k, he can either pay the attention cost $c > 0$ and visit one sender, or take an action with the information gathered so far. After the action is taken, the game ends.

Senders offer *experiments* over their own signal. To avoid signalling, I assume that senders do not observe their signal prior to revealing it through experiments. An experiment is a conditional distribution over messages m from the Polish space M. The message space M is equal for all senders and rich enough to contain all information about $\boldsymbol{x} = (x_1, \cdots, x_I)$, i.e. $X \subset M$. At the beginning of round k, each sender i simultaneously announces $\lambda_{i,k} : X_i \times \mathcal{B}(M) \to [0, 1]$, a (regular) conditional probability such that $\lambda_{i,k}(\cdot, W)$ is measurable for all $W \in \mathcal{B}(M)$, and $\lambda_{i,k}(x_i, \cdot)$ is a probability measure given any signal $x_i \in X_i$. The set of possible experiments for sender i is denoted by Λ_i. Senders compete for attention. In each round that sender i is visited, she receives utility normalised to one.

1.2.2 Strategies, Payoffs, Equilibrium

First, nature draws the state ω and the signals $\boldsymbol{x} = (x_1, \cdots, x_I)$. At the beginning of each round $k \geq 0$, if the receiver has not taken an action previously, all senders simultaneously offer experiments $\lambda_{i,k}$.

The receiver observes the offers and chooses $d_k \in \{0, 1, \cdots, I\}$, where $d_k = 0$ encodes that he stops and $d_k = i \in \{1, \cdots, I\}$ means that he pays cost $c > 0$ and visits sender i. When the receiver stops, he takes an action $a \in A$, the game ends, and payoffs realise.[9] Visiting sender i implies that he observes $m_{i,k} \in M$ drawn from

[8] A Polish space is a separable and completely metrisable space. This ensures the existence of the conditional probability measures used below.

[9] For completeness, assume that $d_k = 0 \Rightarrow d_{k+1} = 0$.

the distribution $\lambda_{i,k}$ and the game continues to the next round.

A public **history** of the game is

$$h^k = \left(\left((\lambda_{i,1})_{i=1}^I, d_1, m_{d_1,1}\right), \cdots, \left((\lambda_{i,k})_{i=1}^I, d_k, m_{d_k,k}\right)\right) \qquad \text{for } k \geq 0,$$

with initial history $h^{-1} = \emptyset$. All players observe all past offers, the receiver's choices and the message of the chosen sender. Hence, all senders observe the information revealed by their competitors.[10] Denote by $\mathcal{H}^k$ the set of round-k histories.

A pure **strategy** for the receiver is a collection of maps $\left(\sigma_k^R\right)_{k\geq 0}$ with

$$\sigma_k^R : \mathcal{H}^{k-1} \times \left(\bigtimes_i \Lambda_i\right) \to \{0, 1, \cdots, I\}.$$

Likewise, for each sender $i \in \{1, \cdots, I\}$, a pure strategy is a collection $(\sigma_k^i)_{k\geq 0}$ with

$$\sigma_k^i : \mathcal{H}^{k-1} \to \Lambda_i.$$

The receiver's final **payoff** has two components. First, he gets utility $u(a, \omega)$ when stopping with action a if the state is ω. Second, there are attention costs that depend on how many times the receiver visits a sender before taking action. Each visit costs $c > 0$, so that the receiver's final payoff will be

$$u(a, \omega) - c \cdot \sum_{k\geq 0} \mathbb{1}_{\{d_k \neq 0\}}.$$

Senders maximise the attention they attract. Each visit gives utility normalised to 1. Sender i's final payoff is then

$$\sum_{k\geq 0} \mathbb{1}_{\{d_k = i\}}.$$

The **solution concept** is a Perfect Bayesian Equilibrium, with the additional requirement that beliefs are only updated with Bayes' rule according to the chosen experiment. This requirement ensures the 'no-signalling-what-you-don't-know' property (see Fudenberg and Tirole, 1991), whereby offers do not reveal information the senders do not hold. It also implies that experiments from off-path offers would be interpreted correctly if they were accepted.

[10] It is plausible that the senders can also visit their competitors, given that experiments are offered publicly. With Gaussian information considered in Section 1.5, I show how this assumption can be relaxed.

1.2.3 Examples

Before moving to the analysis, I give two specific setups encompassed by the model presented above. Applications for which these setups or a close variant have been used in the literature are mentioned in square brackets. The News Markets application is considered more carefully in Section 1.5.

Example 1: Gaussian Information and Quadratic Loss. [Global games: Morris and Shin (2002), Bergemann and Morris (2013), Angeletos and Pavan (2007). Social learning: Vives (1996). News markets: Chen and Suen (2019), Galperti and Trevino (2018)]

The receiver wants to learn the state of the world $\omega \sim \mathcal{N}(0, 1/p_0)$ as precisely as possible. His utility from the action is $u(a, \omega) = -(\omega - a)^2$, so that the expected stopping utility is minus the conditional variance $Var(\omega|\xi)$ given the current information.[11] Each sender $i \in \{1, \cdots, I\}$ is endowed with a conditionally independent signal $x_i \sim \mathcal{N}(\omega, 1/p_i)$, where $p_i > 0$ is sender i's *precision level.*

Example 2: Additive Attributes. [Consumer search: Wolinsky (1986), Choi et al. (2018), Ke and Lin (2020). Advertising: Anderson and Renault (2009), Sun (2011)]

Let the receiver be a consumer who considers buying one of two objects, $A = \{1, 2\}$. The (net) utility of the two objects is $\omega = (\omega_1, \omega_2) \in \mathbb{R}^2$, where each object's utility is determined by a common and an idiosyncratic attribute as follows:

$$\omega_i = Y + \gamma_i \qquad\qquad \text{for } i \in \{1, 2\},$$

where Y is the common component distributed according to F on $[\underline{Y}, \bar{Y}]$ and γ_i are distributed independently according to G_i on $[\underline{\gamma}, \bar{\gamma}]$. Each sender holds information about one of the options in the form of a noisy signal of the total utility but is unable to distinguish between the common and the idiosyncratic component: $x_1 = \omega_1 + \epsilon_1$ and $x_2 = \omega_2 + \epsilon_2$ with $\epsilon_i \overset{\text{iid}}{\sim} \mathcal{N}(0, 1)$. Senders in my model are indifferent about the receiver's actions. This is the case if the sender does not sell the product by herself, as with car or technology magazines on- and offline.

1.3 Updating and the Value of Information

This section introduces notation that will be used extensively in the remainder of the analysis. To determine the optimal action and compute the expected utility from stopping, the receiver has to form a belief about ω. The messages deliver

[11] The fact that u is not bounded from below does not create problems here since the stopping utility at the prior is equal to $-\frac{1}{p_0} > -\infty$.

information about signals $\boldsymbol{x} = (x_1, \cdots, x_I)$. Recall that the joint distribution over the state and signals is $\tilde{\mu} \in \Delta(\Omega \times X)$. I denote its marginal with respect to $\boldsymbol{x}$, that is, the unconditional prior distribution of signals, by $\xi^0 \in \Delta(X)$. It will be convenient to work with the posterior signal-belief ξ in the rest of the paper.

The receiver's *stopping utility* with belief ξ, i.e. the expected utility from the optimal action, given that he currently holds posterior ξ, is

$$U(\xi) \equiv \max_{a \in A} \mathbb{E}_{\omega \sim \mu(\xi)} \left[\, u(a, \omega) \, \right],$$

where $\mu(\xi)$ denotes the belief about the state given that the belief about the signals is ξ. The belief μ is not used in the analysis. The formula is given in Appendix 1.A.1 together with further details on this section.

At the end of round k, the belief ξ^k is updated to ξ^{k+1} by Bayes' rule after observing message m_k resulting from the selected experiment λ_k. Denote the *updating rule* by ξ' such that

$$\xi^{k+1} = \xi'(\xi^k, m).$$

Note that the notation suppresses the chosen experiment in the updating rule. Appendix 1.A.1 contains the updating rule in detail and shows that it is well defined. I denote by $\xi'(\xi^0, \boldsymbol{x}_{-i})$ the belief that results if the signals of all senders different from i are known and nothing has been revealed about x_i.

With this, define the *value* of offer λ_i at current belief ξ as

$$v\left(\lambda_i | \xi\right) \equiv \mathbb{E}_{x_i \sim \xi} \left[\mathbb{E}_{m \sim \lambda_i(x_i, \cdot)} \left[U\left(\xi'(\xi, m)\right) \right] \right] - U(\xi).$$

The value is defined as the expected difference between the stopping utilities with and without the additional information from λ_i. Note that $v \geq 0$ always.

For the special case in which sender i's experiment reveals her exact signal, let

$$\bar{v}_i(\xi) \equiv v\left(\delta_{\{x_i\}} | \xi\right)$$

denote the value of all her information given belief ξ. Here, the experiment that reveals i's signal precisely is denoted by the Dirac measure $\lambda_i(x_i, \cdot) = \delta_{\{x_i\}}(\cdot)$.

I assume that attention can be split finely enough to make every sender's information worth one unit of attention, independent of the realisation of her opponents' signals.

Assumption 1. *For all $i \in \{1, \cdots, I\}$ and for all $\boldsymbol{x}_{-i}$ in the support of $\xi^0(X_i, \cdot)$:*

$$\bar{v}_i\left(\xi'(\xi^0, \boldsymbol{x}_{-i})\right) > c. \tag{A1}$$

This condition ensures that for any realisation of the other senders' signals, even if the receiver knows these exactly, sender i still has enough information to attract at least one visit. In particular, condition (A1) implies that no sender has perfect information. Given the Bertrand competition, if at least two senders had perfect information, all senders would offer all information in the first round, and the receiver would become perfectly informed after one visit.[12]

1.4 Equilibrium

This section identifies a simple class of information-transmission processes that is sufficient to achieve the unique equilibrium payoffs in the case of a single sender and the receiver-preferred equilibrium payoff with multiple senders, in which all information is transmitted in the shortest amount of time possible.

1.4.1 Single Sender

Consider the case of a single sender, $I = 1$. What is the maximal expected attention cost the receiver is willing to pay for the sender's information? It is equal to the difference between the stopping utility with no information and the expected stopping utility with all information. This is precisely $\bar{v}_1(\xi^0)$. As each visit requires a cost of c, the maximal expected number of visits the sender can attract is

$$\frac{\bar{v}_1\left(\xi^0\right)}{c}.$$

If the sender could commit across rounds, the simplest strategy to implement this outcome would require $\frac{\bar{v}_1(\xi^0)}{c} - 1$ visits from the receiver at which no information is revealed, and then all information would be revealed at the last visit.[13] However, the sender lacks the intertemporal commitment to credibly promise all information in the last round. She may, for example, repeat the round-0 strategy.

A simple sender strategy to overcome the non-commitment issue and to deal with potential integer problems is to offer revealing x_1 with probability $\lambda^* \in [0,1]$ and revealing no information with probability $1 - \lambda^*$.[14] Offers in this class are denoted as *All-or-Nothing* (AoN) offers. To give a simple example of an AoN offer, assume that the sender's signal is the result of a coin flip, $x_1 \in \{0,1\}$, with $\xi^0 = 0.6$ prior probability that $x_1 = 1$. Let the receiver's utility be 1 if he guesses the signal

[12] See Section 1.5 for the equivalent of (A1) for Example 1. For Example 2, one sufficient condition for (A1) is that the noise term of each sender has sufficient variance.

[13] This intuitive argument neglects non-divisibilities that make this potential strategy suboptimal as it could only achieve integer amounts of visits.

[14] I abuse the notation by letting λ^* denote a probability in [0,1], while λ generally denotes distributions over messages. Formally, the AoN experiment is represented by the conditional distribution $\lambda(x_1, \cdot) = \lambda^* \delta_{\{x_1\}} + (1 - \lambda^*)\delta_{\{m\}}$ for an arbitrary message $m \in M \backslash X_i$ that conveys no information.

correctly and 0 otherwise. That is, $\omega = x_1 \in \{0,1\}$, $A = \{0,1\}$, and $u(a,\omega) = \mathbb{1}_{\{a=\omega\}}$.[15]

Figure 1.1: AoN experiment with binary signal

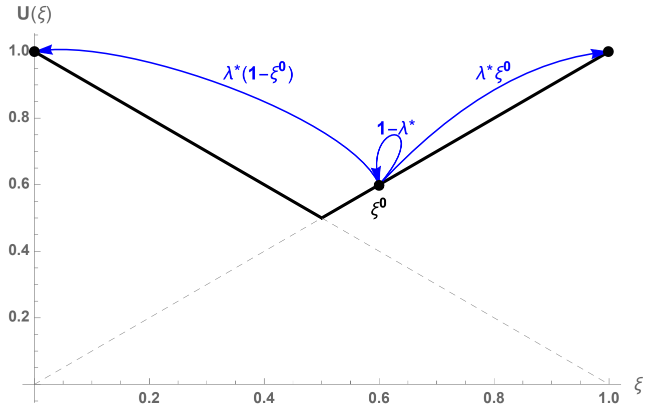


Figure 1.1 shows the receiver's expected utility as a function of the belief as a solid black line. The three arrows indicate the possible jumps in belief induced by the AoN experiment. As the middle arrow shows, with probability $1-\lambda^*$, the experiment reveals no information and the belief remains unchanged. With probability λ^*, the sender's signal is revealed perfectly so that with probability $\lambda^*(1-\xi^0)$, the belief jumps left to 0, and with probability $\lambda^*\xi^0$, the belief jumps right to 1.

The following result shows that an equilibrium in AoN strategies generally exists in the monopoly game. Equilibrium payoffs are unique.

Lemma 1. *Let $I = 1$. There is an AoN equilibrium in which, in each round, the sender offers AoN probability*

$$\lambda_1^*(\xi^0) = \frac{c}{\bar{v}_1(\xi^0)}.$$

The receiver accepts every round until x_1 is revealed. In any equilibrium of the monopoly game the expected payoffs are $\frac{\bar{v}_1(\xi^0)}{c}$ for the sender and $U(\xi^0)$ for the receiver.

Note that assumption (A1) ensures that $\lambda_1^*(\xi^0) < 1$. If the sender's strategy prescribes AoN offers until all information is transmitted, the receiver's continuation payoff in the event of no revelation (which happens with probability $1 - \lambda_1^*(\xi^0)$) remains at his initial payoff. Hence, the AoN probability $\lambda_1^*(\xi^0)$ that makes the receiver indifferent between taking action immediately and accepting the offer, has

[15] Here, state and signal are identical. For this section, the general model could equivalently be specified with $\omega = \boldsymbol{x}$. However, in Section 1.5, signals are chosen endogenously by the senders, so that modelling the payoff-relevant state separately allows keeping the endogenous signals and the exogenous state distribution apart.

to satisfy

$$U\left(\xi^0\right) = -c + \lambda_1^*(\xi^0)\, \mathbb{E}_{x_1\sim\xi^0}\left[U\left(\xi'(\xi^0, x_1)\right)\right] + (1-\lambda_1^*(\xi^0))U\left(\xi^0\right).$$

Accepting the offer creates attention cost c. With probability $\lambda_1^*(\xi^0)$, the receiver learns x_1 and stops with utility $U(\xi'(\xi^0, x_1)))$. With probability $1 - \lambda_1^*(\xi^0)$, the receiver learns no information, which gives utility $U(\xi^0)$ as the sender will keep him indifferent in the following round again.

The offer $\lambda_1^*(\xi^0)$ is accepted by the receiver in every round until the information is eventually revealed. The number of rounds until revelation follows a geometric distribution with parameter $\lambda_1^*(\xi^0)$, so that the expected number of rounds is $\frac{1}{\lambda_1^*(\xi^0)}$. As the receiver is indifferent between accepting and stopping in every round, it should not be a surprise that solving the above indifference condition for λ^* gives

$$\lambda_1^*(\xi^0) = \frac{c}{\bar{v}_1\left(\xi^0\right)}.$$

The expected attention is precisely the upper bound the receiver is willing to spend.

Depending on the information structure, there may be other strategies that resolve the sender's non-commitment and attract the maximal amount of attention.[16] The attractiveness of the AoN strategy lies in the fact that it works for general information structures and in its simplicity. It leads to a stationary information-arrival process. Furthermore, the lack of inter-temporal commitment requires that any experiment delivers, in expectation, a strictly positive increase in the stopping utility. In Section 1.6.2, I discuss in detail that this strict increase requires beliefs to jump with positive probability whenever the action set is finite and how – even in the limit as the period length shrinks – the information process necessarily features a Poisson-jump component. Thereby, this identifies an additional driver of Poisson information, stemming from the lack of commitment, rather than risk preferences induced by discounting (see Zhong, 2019). The above results are robust to discounting; the analysis remains almost unchanged when the receiver and the sender share a common discount factor.

1.4.2 Multiple Senders

In the general case with $I \geq 2$ senders, equilibrium payoffs are no longer unique. The subsequent analysis focuses on receiver-preferred equilibria. This selection best

[16]The sender could reveal some information every round, successively increasing the receiver's stopping utility to commit herself to offer even more information in the following round. Appendix 1.A.2 includes an example of such a process when information is normally distributed. The special feature of the Gaussian distributions allows the sender to achieve information transmission in a deterministic number of visits (modulo integer problems).

captures the tradeoff between the amount and the speed of information transmission since information has no instrumental value for the senders. The welfare-maximising equilibrium crucially depends on the normalisation of the value the senders derive from each visit. In particular, if $c = 1$, maximising welfare is equivalent to maximising the amount of information transmitted as visits from the receiver to any sender have no impact on welfare. Whenever $c \geq 1$, the receiver-preferred equilibrium is also welfare-maximising.

With multiple senders, there are informational externalities that may impede information transmission. For illustration, consider the following example with the detailed argument presented in Section 1.6.3. Suppose there are two senders. Each sender's signal is an independent, fair coin flip. The receiver has to guess whether the two coins match or not. For this decision problem, the signals form complements. Each signal is valuable only in conjunction with the other. Since senders cannot commit across rounds, complements cause a hold-up problem: after one sender has revealed her information, the following sender would require maximal attention, keeping the receiver at the current stopping utility with one signal only. Anticipating this, the receiver is not willing to spend any attention for the first signal given that it delivers no value on its own.

To rule out this class of problems and ensure information transmission, I introduce the following condition.

Definition 1. *The senders' signals are substitutes if, for all i and for all beliefs ξ with $supp(\xi) \subseteq supp(\xi^0)$:*

$$\bar{v}_i(\xi) \geq \mathbb{E}_{\boldsymbol{x}_{-i} \sim \xi}\left[\bar{v}_i\left(\xi'(\xi, \boldsymbol{x}_{-i})\right)\right]. \tag{SU}$$

Signals are substitutes if the current value of x_i at belief ξ is greater than the expected value after knowing all other senders' signals.[17] That sender i's information is more valuable the less is known from her competitors is consistent with many applications. This is especially the case when senders report on a single issue or, as in Example 2, when the signals allow inference about a common component that affects all options. In both examples above, signals are substitutes.

In equilibrium, competing senders make offers that make them indifferent between being accepted or rejected. Constructing an AoN equilibrium requires determining the maximal AoN probability a competing sender is willing to offer. Consider the situation in which all senders but i have revealed their signals, and sender i has revealed no information at all. That is, $\boldsymbol{x}_{-i}$ is known and the belief is $\xi'(\xi^0, \boldsymbol{x}_{-i})$.

[17] Börgers et al. (2013) introduce notions of substitutes and complements for a pair of signals. Viewing x_i and $\boldsymbol{x}_{-i}$ as two signals, (SU) corresponds to the notion of substitutability in Börgers et al. (2013) for given ξ and restricted to the specific decision problem considered here. Their requirement is independent of the decision problem and therefore stronger.

Sender i can extract maximal attention from the receiver. The receiver is willing to visit her

$$\frac{\bar{v}_i\left(\xi'(\xi^0, \boldsymbol{x}_{-i})\right)}{c}$$

rounds to learn x_i.

As will be shown below, a lower bound on the attention sender i can extract is given by waiting until all competitors have revealed their information and offering AoN probability

$$\lambda_i^*\left(\xi'(\xi^0, \boldsymbol{x}_{-i})\right) = \frac{c}{\bar{v}_i\left(\xi'(\xi^0, \boldsymbol{x}_{-i})\right)}.$$

However, this value depends on the realisations of $\boldsymbol{x}_{-i}$, so that its expectation – taken over all competitors' signals given the current information – changes over time. Suppose the AoN probability offered by each sender is such that her expected payoff is precisely the expectation of the outside option mentioned above, assuming that this offer was repeatedly accepted until revelation. The following result shows that, if signals are substitutes, these strategies form an equilibrium. In addition, this equilibrium attains full information transmission in the shortest possible time among all equilibria, making it receiver-preferred.

Theorem 1. *If senders' signals are substitutes, there is an equilibrium with the following strategies. At belief ξ, senders whose information has not been revealed make AoN offers with probability*

$$\lambda_i^*(\xi) = \frac{c}{\mathbb{E}_{\boldsymbol{x}_{-i}\sim\xi}\left[\bar{v}_i\left(\xi'(\xi^0, \boldsymbol{x}_{-i})\right)\right]}. \tag{1.1}$$

The receiver is indifferent between visiting any of the senders whose information has not been revealed and visits them in arbitrary order until all information is transmitted. This equilibrium is receiver-preferred.

Proof. Note that the result characterises a class of equilibria rather than a single equilibrium as the receiver's behaviour is not fixed. By (A1), we have that $\lambda_i^*(\xi) < 1$ for all i whose information has not been revealed. The proof is organised in three claims:

Claim 1. *Fix any strategies by senders $\neq i$ and assume the receiver is playing a best response. Let the current belief be ξ and assume sender i has not revealed any information. Then, playing the AoN strategy from the theorem secures sender i an expected payoff of $\frac{1}{\lambda_i^*(\xi)}$.*

Proof of Claim 1. First, we show that the AoN strategy ensures that the receiver will not stop without observing sender i's information. Formally, for all beliefs ξ,

$$-c + \lambda_i^*(\xi) E_{x_i \sim \xi}\left[V_R\left(\xi'(\xi, x_i)\right)\right] + (1 - \lambda_i^*(\xi)) V_R(\xi) \geq U(\xi).$$

Here, V_R denotes the receiver's continuation value (suppressing history and strategy). Clearly, $V_R(\xi) \geq U(\xi)$, as the receiver always has the option to stop. For the above inequality to hold, substituting and rearranging gives that it is sufficient to show that

$$E_{x_i \sim \xi}\left[U\left(\xi'(\xi, x_i)\right)\right] - U(\xi) \geq \frac{c}{\lambda_i^*(\xi)}.$$

The left-hand side of this inequality is the definition of $\bar{v}_i(\xi)$. Replacing λ_i^* on the right-hand side with (1.1) shows that this inequality is equivalent to the definition of substitutes in (SU).

Second, we show that the expected payoff for sender i from using the AoN strategy is exactly $\frac{1}{\lambda_i^*(\xi)}$. To illustrate this concisely, the remainder of the argument for Claim 1 considers Markov strategies, so that the belief ξ determines the senders' payoffs. This restriction is not necessary for the result and a detailed argument without it is included in Appendix 1.A.1. Observe that i's valuation satisfies:

$$V_i(\xi) = \begin{cases} 1 + \lambda_i^*(\xi)0 + (1 - \lambda_i^*(\xi)) V_i(\xi) & \text{if } i \text{ is chosen} \\ 0 + \mathbb{E}_{x_j \sim \xi}\left[\mathbb{E}_{m_j \sim \lambda_j(x_j, \cdot)}\left[V_i(\xi'(\xi, m_j))\right]\right] & \text{if } j \neq i \text{ is chosen.} \end{cases}$$

In the first line, i is visited and her continuation value is 0 if her information is revealed and remains unchanged if no information is given out. The value $\frac{1}{\lambda_i^*(\xi)}$ follows immediately from re-arranging. In the second line, depending on the realisation of m_j and the receiver's choice in the following round, the value $V_i(\xi'(\xi, m_j))$ is either $\frac{1}{\lambda_i^*(\xi'(\xi, m_j))}$ if sender i is chosen in that round, or

$$V_i(\xi'(\xi, m_j)) = \mathbb{E}_{x_\ell \sim \xi'(\xi, m_j)}\left[\mathbb{E}_{m_\ell \sim \lambda_\ell(x_\ell, \cdot)}\left[V_i(\xi'(\xi'(\xi, m_j), m_\ell))\right]\right],$$

if a sender $\ell \neq i$ is chosen. As $c > 0$, there can be at most finitely many rounds and realisations before sender i is chosen so that, eventually, we arrive at realisations with belief $\hat{\xi}$ and $V_i(\hat{\xi}) = \frac{1}{\lambda_i^*(\hat{\xi})}$.

Since, by definition,

$$\frac{1}{\lambda_i^*(\hat{\xi})} = \mathbb{E}_{\boldsymbol{x}_{-i} \sim \hat{\xi}}\left[\bar{v}_i\left(\xi'(\xi^0, \boldsymbol{x}_{-i})\right)\right] \frac{1}{c},$$

we have that

$$
\begin{aligned}
&\mathbb{E}_{x_j\sim\xi}\left[\mathbb{E}_{m_j\sim\lambda_j(x_j,\cdot)}\left[\frac{1}{\lambda_i^*(\xi'(\xi,m_j))}\right]\right] \\
&= \mathbb{E}_{x_j\sim\xi}\left[\mathbb{E}_{m_j\sim\lambda_j(x_j,\cdot)}\left[\mathbb{E}_{\boldsymbol{x}_{-i}\sim\xi'(\xi,m_j)}\left[\bar{v}_i\left(\xi'(\xi^0,\boldsymbol{x}_{-i})\right)\right]\frac{1}{c}\right]\right] \\
&= \mathbb{E}_{\boldsymbol{x}_{-i}\sim\xi}\left[\bar{v}_i\left(\xi'(\xi^0,\boldsymbol{x}_{-i})\right)\frac{1}{c}\right] = \frac{1}{\lambda_i^*(\xi)}.
\end{aligned}
$$

Therefore, by taking expectations as many times as necessary from the last realisation to the current stage with belief ξ, we get the claimed payoff.

Claim 2. *Let all senders play the AoN strategies from the theorem and assume the receiver is playing a best response. Then, no sender has a profitable deviation.*

Proof of Claim 2. The receiver visits all senders on the equilibrium path by Claim 1. Suppose the on-path belief is ξ and that only information from senders $1,\cdots,j-1$ has been observed. Then, the equilibrium continuation utility of the receiver can be expressed as

$$
V_R(\xi) = \mathbb{E}_{\boldsymbol{x}\sim\xi}\left[U\left(\xi'(\xi^0,\boldsymbol{x})\right)\right] - c\sum_{i=j}^{I} V_i(\xi), \tag{1.2}
$$

where the proof of the last claim showed that $V_i(\xi) = \frac{1}{\lambda_i^*(\xi)}$ for all i.

Suppose now that one sender i' deviates to an offer that gives an expected payoff higher than $\frac{1}{\lambda_{i'}^*(\xi)}$ if accepted. By visiting the remaining, non-deviating senders, the receiver achieves an expected payoff of

$$
\mathbb{E}_{\boldsymbol{x}_{-i'}\sim\xi}\left[U\left(\xi'(\xi^0,\boldsymbol{x}_{-i'})\right)\right] - c\sum_{\substack{i=j\\ i\neq i'}}^{I}\frac{1}{\lambda_i^*(\xi)}. \tag{1.3}
$$

The λ_i^* are chosen such that, at the last sender, the receiver is indifferent between stopping without her information or paying the corresponding attention cost to obtain her information. Hence, (1.2) and (1.3) are equal. Even if the alternative strategy of sender i' would lead to all her information being revealed, the receiver still prefers to reject any offer yielding i' an expected payoff higher than $\frac{1}{\lambda_{i'}^*(\xi)}$.

Claim 3. *The AoN equilibrium achieves the maximal payoff for the receiver among all equilibria.*

Proof of Claim 3. The action is always taken with all information. As the value of information is always positive and, by claim 1, no sender can receive less attention in expectation, the AoN equilibrium is receiver-preferred.

□

Theorem 1 yields a simple computation of the equilibrium payoffs of the receiver and the senders. It suffices to compute the expected residual value of each sender's signal. After the following remark, the next subsection makes use of this to show that, if signals are substitutes, concentrating information on fewer senders slows down transmission.

Remark 1. *Note that the on-path strategies in this equilibrium are Markov. Senders' actions are fully determined by the state, ξ. The receiver's actions are fully determined by the state and the offers made in the current round. The state ξ is not strictly payoff-relevant in that two distinct posterior beliefs over signals may lead to the same belief μ over states and, therefore, to the same stopping utility. One might argue that μ is a more appropriate state variable. However, it is easy to verify that the belief μ captures too little information to determine optimal behaviour in the continuation game: consider a case with two senders who have symmetrically distributed signals. One of the two has revealed this signal to the receiver but the other has not. The same belief μ may be derived from sender 1's signal or sender 2's signal being known, but the value of future offers from senders 1 and 2 depends crucially on this distinction.*

1.4.3 Concentration of Information

As the action is taken after all information is transmitted, the receiver's expected total payoff is

$$\mathbb{E}_{\boldsymbol{x}\sim\xi^0}\left[U\left(\xi'(\xi^0,\boldsymbol{x})\right)\right] - c\sum_{i=1}^{I}\frac{1}{\lambda_i^*(\xi^0)}.$$

Each sender's attention is proportional to the residual value of her information so that the receiver's total payoff is equal to

$$\mathbb{E}_{\boldsymbol{x}\sim\xi^0}\left[U\left(\xi'(\xi^0,\boldsymbol{x})\right)\right] - \sum_{i=1}^{I}\mathbb{E}_{\boldsymbol{x}_{-i}\sim\xi^0}\left[\mathbb{E}_{x_i\sim\xi'(\xi^0,\boldsymbol{x}_{-i})}\left[U\left(\xi'(\xi^0,\boldsymbol{x})\right)\right] - U\left(\xi'(\xi^0,\boldsymbol{x}_{-i})\right)\right].$$

Consider increasing the concentration of information by merging the signals of senders $i = 1$ and $i = 2$ into a single signal $x_{1,2} = (x_1, x_2)$ held by sender 2. Sender 1 has no information and is excluded from the game.

This decreases the speed of information transmission. To see this, consider the receiver's utility after the concentration. The first term remains the same as the overall information has not changed. The attention required by senders $i \in \{3, \cdots, I\}$ also remains unaffected. The change comes from the attention required for signal

$x_{1,2}$, which is now equal to

$$\mathbb{E}_{\boldsymbol{x}\sim\xi^0}\left[U\left(\xi'(\xi^0,\boldsymbol{x})\right)\right]-\mathbb{E}_{\boldsymbol{x}_{-\{1,2\}}\sim\xi^0}\left[U\left(\xi'(\xi^0,\boldsymbol{x}_{-\{1,2\}})\right)\right]. \tag{1.4}$$

Before the concentration, observing x_1 and x_2 required the attention of

$$\mathbb{E}_{\boldsymbol{x}\sim\xi^0}\left[U\left(\xi'(\xi^0,\boldsymbol{x})\right)\right]-\mathbb{E}_{\boldsymbol{x}_{-1}\sim\xi^0}\left[U\left(\xi'(\xi^0,\boldsymbol{x}_{-1})\right)\right]+\mathbb{E}_{\boldsymbol{x}\sim\xi^0}\left[U\left(\xi'(\xi^0,\boldsymbol{x})\right)\right] \tag{1.5}$$
$$-\,\mathbb{E}_{\boldsymbol{x}_{-2}\sim\xi^0}\left[U\left(\xi'(\xi^0,\boldsymbol{x}_{-2})\right)\right].$$

To see that the cost after concentrating the information in (1.4) is greater than the cost before in (1.5), consider the difference:

$$\mathbb{E}_{\boldsymbol{x}_{-1}\sim\xi^0}\left[U\left(\xi'(\xi^0,\boldsymbol{x}_{-1})\right)\right]+\mathbb{E}_{\boldsymbol{x}_{-2}\sim\xi^0}\left[U\left(\xi'(\xi^0,\boldsymbol{x}_{-2})\right)\right]$$
$$-\mathbb{E}_{\boldsymbol{x}_{-\{1,2\}}\sim\xi^0}\left[U\left(\xi'(\xi^0,\boldsymbol{x}_{-\{1,2\}})\right)\right]-\mathbb{E}_{\boldsymbol{x}\sim\xi^0}\left[U\left(\xi'(\xi^0,\boldsymbol{x})\right)\right].$$

This can be rearranged to

$$\mathbb{E}_{\boldsymbol{x}_{-\{1,2\}}\sim\xi^0}\left[\bar{v}_1(\xi'(\xi^0,\boldsymbol{x}_{-\{1,2\}}))\right]-\mathbb{E}_{\boldsymbol{x}_{-1}\sim\xi^0}\left[\bar{v}_1(\xi'(\xi^0,\boldsymbol{x}_{-1}))\right],$$

which is positive since signals are substitutes. Hence, concentrating the same amount of information on fewer senders slows down information transmission and hurts the receiver.

1.5 News Markets

This section applies the main results to the specification in Example 1 to derive and interpret further comparative statics and extend the model by endogenous information acquisition. Variants of this Gaussian setting have been applied to study various aspects of media markets in Galperti and Trevino (2018), Chen and Suen (2019), and others.

The receiver wants to be informed about the state of the world $\omega\sim\mathcal{N}(0,1/p_0)$. He wants to match the state with his action a and gets utility $u(a,\omega)=-(a-\omega)^2$. Each newspaper i holds some information about the state represented by a signal that is independent conditional on the state: $x_i\sim\mathcal{N}(\omega,1/p_i)$ where $p_i>0$ is called i's *precision level.* The receiver's optimal action at belief ξ is $a^*(\xi)=\mathbb{E}_{\omega\sim\mu(\xi)}[\omega]$, and the expected utility from stopping with belief ξ is $-\mathbb{E}_{\omega\sim\mu(\xi)}\left[(a^*(\xi)-\omega)^2\right]=-Var\left(\omega\mid\xi\right)$. Hence, the receiver's stopping utility at prior information is $-\frac{1}{p_0}$. The reduction in variance caused by any sender's signal is independent of the realisation of her own or her opponents' signal. In particular, if the receiver knows the signals of all senders, his stopping utility is $-\frac{1}{p_0+p_1+\ldots+p_I}$. Precision increases linearly.

1.5.1 Exogenous Precision

To characterise the AoN equilibrium analogous to Theorem 1, define by

$$P \equiv p_0 + \sum_{i=1}^{I} p_i,$$

the precision level of all senders plus p_0, the precision of the state distribution. Let $P_{-i} = P - p_i$ denote the total precision without sender i. In this case, assumption (A1) boils down to the requirement that, for all i:

$$\frac{1}{P_{-i}} - \frac{1}{P_{-i} + p_i} > c,$$

so that the residual value of sender i's information exceeds the attention cost c.

Corollary 1. *There is an AoN equilibrium analogous to Theorem 1 in which each sender i offers AoN probability*

$$\lambda_i^* = c\frac{P_{-i}(P_{-i} + p_i)}{p_i}$$

in every round until her signal is revealed.

In this AoN equilibrium, sender i expects to attract the total attention of

$$\frac{1}{\lambda_i^*} = \frac{p_i}{cP_{-i}(P_{-i} + p_i)}.$$

With higher precision, she can attract more attention. A higher precision of her competitors' signals or of the initial distribution both lead to a higher P_{-i} and decrease sender i's expected attention.

The receiver's final payoff is

$$-\frac{1}{P} - c\sum_{i=1}^{I} \frac{p_i}{cP_{-i}(P_{-i} + p_i)}.$$

Hence, fixing the total precision level P, the reader is better off, the smaller is

$$\sum_{i=1}^{I} \frac{p_i}{P - p_i} = \sum_{i=1}^{I} \left(\frac{P}{P - p_i} - 1 \right).$$

The highest utility the receiver can get is trivially achieved at maximal prior precision with $p_0 = P$ and $p_i = 0$ for all $i \geq 1$. If we fix P and p_0, does the receiver prefer the remaining precision to be distributed evenly among all newspapers or to be skewed with some papers holding a lot and others holding very little information? The fraction on the right side of the equality is convex in p_i. The receiver prefers a uniform distribution of precision levels over senders, that is, $p_i = \frac{P-p_0}{I}$ for all $i \in \{1, \cdots, I\}$.

1.5.2 Information Acquisition

In practice, the precision of a newspaper's information is endogenously determined by its investigation process and editorial policies. One crucial factor that affects precision is the time at which a story is reported. Investigating a newsworthy issue features a natural tradeoff between checking further sources more carefully and running a story as early as possible.[18]

This subsection considers an *investigation race* between two newspapers to examine how this time tradeoff affects precision levels. Each paper's precision is determined by the time elapsed until it starts reporting the story. The following results show that the investigation race leads to specialisation of the two papers into an early reporter with lower precision and a late reporter with higher precision. This is the case even if their productivity levels, the increase in precision per investigated time, are identical. More than that, when the precision levels are unequal, the investigation race exacerbates the inequality: the more productive newspaper will deepen its advantage by investigating longer than the less productive competitor. Further comparative statics offered below show that increasing initial public precision may lead to a decrease of total final precision.

To allow for cleaner exposition, the following results are presented in terms of a continuous-time game in which newspaper i's precision level is $k\rho_i$ after market entry at time k. The increase in precision per instant, ρ_i, can be interpreted as the investigation productivity of newspaper i. That is, investigating from time 0 until some time $k \geq 0$ results in a signal $x_i \sim \mathcal{N}(\omega, \frac{1}{k\rho_i})$.[19] Appendix 1.A.2 presents the discrete-time game underlying this subsection. The outcomes presented here are to be interpreted as equilibrium results in the discrete-time game, considering arbitrarily short periods. I assume that the receiver incurs attention costs only after the first sender entered the market. One interpretation is that the issue at hand only becomes eminent for the receiver after the first piece of news is offered. Furthermore, I assume senders are productive enough so that their investigation is initially worthwhile from the receiver's perspective: the marginal increase in utility, $\frac{\partial}{\partial k}\frac{-1}{p_0+\rho_i k}$, at $k = 0$ is higher than the marginal cost, or, equivalently:

[18] See the paper 'The thirst to be first' by Lewis and Cushion (2009) for a discussion of the importance of breaking news earlier than competitors.

[19] This arises for example if we assume that paper i observes a Brownian motion with drift ω and instantaneous variance $\frac{1}{\rho_i}$.

Assumption 2. *For both newspapers* $i = 1, 2$*:*

$$\rho_i > p_0^2 c. \tag{A2}$$

I focus on pure-strategy equilibria and restrict attention to equilibria in AoN strategies such that, after both newspapers entered the market, they play the AoN equilibrium from the previous subsection. In what follows, such equilibria are called *pure AoN equilibria.* This restriction rules out collusive equilibria in which the equilibrium selection after the second paper enters is used to punish or reward specific entry choices. See Appendix 1.A.2 for a discussion of other equilibria.

The first result for the entry game states that there cannot be a pure AoN equilibrium in which both newspapers enter the market at the same time. Consequently, I will refer to the first paper to enter the market as the *leader* and to the second paper as the *follower.*

Lemma 2. *In any pure AoN equilibrium of the investigation race, senders enter at different times. Suppose the leader,* $i = \ell$*, enters at time* k_ℓ*. Then,*

i) the follower, $i = f$*, enters only after the leader has revealed all information.*

ii) the leader's expected payoff is

$$\mathbb{E}[k_f] - k_\ell = \frac{1}{c}\left(\frac{-1}{p_0 + k_\ell \rho_\ell} - \frac{-1}{p_0}\right).$$

iii) if the follower enters at k_f*, her expected payoff is*

$$\frac{1}{c}\left(\frac{-1}{p_0 + k_\ell \rho_\ell + k_f \rho_f} - \frac{-1}{p_0 + k_\ell \rho_\ell}\right).$$

Lemma 2 states further that (i) the follower will enter the market only once the leader has no private information. At this point, not entering would induce the receiver to stop. As long as the leader has enough private information to keep the receiver engaged, the follower prefers to increase her precision and enter later. More precision gives the follower a higher payoff in the AoN equilibrium after she enters. In turn, the leader will not risk the receiver stopping as long as she has enough private information. Item (ii) says that the leader keeps the receiver indifferent between stopping at prior information (with utility $\frac{-1}{p_0}$) and observing the leader's information.[20] Similarly for item (iii), the follower is a monopolist once she enters the market[21] and keeps the receiver indifferent between stopping at the current

[20] By item (i), there are no competing offers from the follower.
[21] The leader's information was fully revealed before.

information (utility $\frac{-1}{p_0+k_\ell \rho_\ell}$) and stopping with both senders' information (utility $\frac{-1}{p_0+k_\ell \rho_\ell + k_f \rho_f}$).

The two papers separate into different editorial policies. The leader starts informing after checking fewer sources, and the follower investigates as long as possible to deliver more in-depth information.

The payoffs in Lemma 2 pin down the expected payoffs in the investigation race as a function of the leader's identity and her entry time k_ℓ. Let $L_i(k_\ell)$ be paper i's payoff if it enters as the leader ($\ell = i$) at time k_ℓ. Let $F_i(k_\ell)$ be paper i's payoff if it becomes the follower as paper $\ell \neq i$ enters at time k_ℓ. The following result collects the properties of the functions L_i and F_i that allow characterising the unique pure AoN equilibrium of the investigation race.

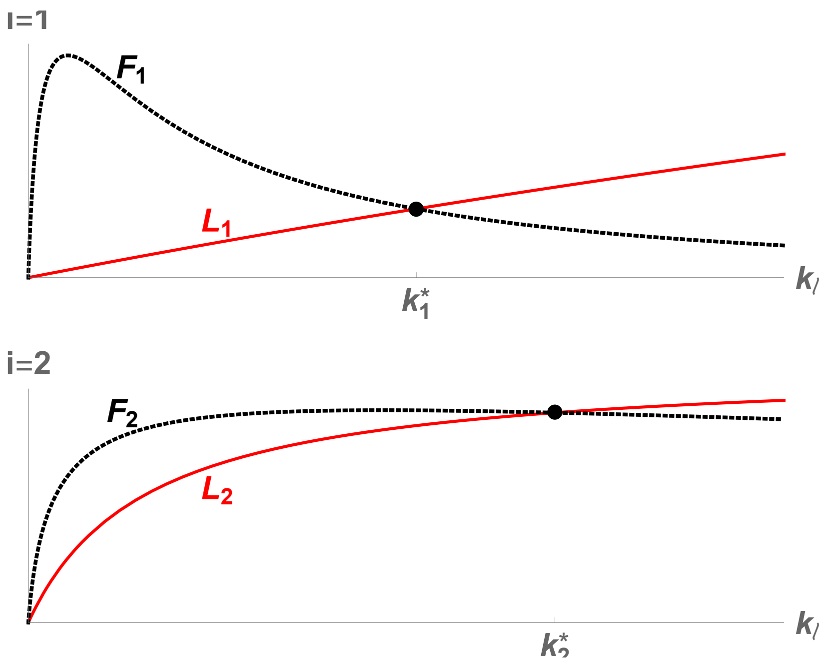
Theorem 2. *For both newspapers $i = 1, 2$;*

- *the leader's payoff $L_i(k_\ell)$ is strictly increasing for all $k_\ell \geq 0$.*
- *there is a time $k_i^* > 0$ with the property that, for all $k_\ell \leq k_i^*$, we have $F_i(k_\ell) \geq L_i(k_\ell)$, and for all $k_\ell > k_i^*$, we have $F_i(k_\ell) < L_i(k_\ell)$.*
- *$k_1^* < k_2^*$ if and only if $\rho_1 < \rho_2$.*

In the unique pure AoN equilibrium of the investigation race, the less productive paper starts reporting first at time $k^* = \max\{k_1^*, k_2^*\}$.

To gain intuition for this result, consider Figure 1.2, which depicts the case in which $\rho_2 > \rho_1$. After $k^* = \max\{k_1^*, k_2^*\} = k_2^*$, both papers strictly prefer to enter as the leader. By continuity, for any potential leader entry time later than k^*, the follower prefers to undercut slightly. For paper i, entering the market as the leader at any $k < k_i^*$ is dominated by entering at k_i^*: if the competitor does not enter before, this is due to the monotonicity of L_i, if the competitor does consider entry before, this is due to $F_i > L_i$. Paper 2 will not enter as the leader before k_2^*, the time at which she is indifferent between entering and becoming the leader or becoming the follower by 1's entry. If 2 does not stop at k_2^* (but at any time strictly later), the best response of paper 1, is to enter as the leader at k_2^*.

Figure 1.2: Leader and Follower Payoffs



Theorem 2 shows that the investigation race presented in this model exacerbates the inequality in precision levels. The more productive paper investigates longer. This resonates with a news cycle in which the 'yellow press' paper first runs a news story with less careful fact-checking, and a more investigative newspaper informs the receiver later but more precisely.

With Theorem 2, we can do comparative statics on the total information discovered in equilibrium. Assume from now on that $\rho_2 \geq \rho_1$, so that newspaper 1 is the first one to enter the market at $k^* = k_2^*$. Then, the expected utility of the receiver from the action is $\mathbb{E}_{k_2}\left[\frac{-1}{p_0+k^*\rho_1+k_2\rho_2}\right]$. This gives a measure for the total information obtained in this game. The following result considers how it changes in parameter values.

Lemma 3. *Holding all other parameters fixed, in the investigation race equilibrium,* $\mathbb{E}_{k_2}\left[\frac{-1}{p_0+k^*\rho_1+k_2\rho_2}\right]$ *is*

i) decreasing in c, and

ii) decreasing in p_0 for all $p_0 \in [0, \underline{p}]$ with $\underline{p} > 0$.

Point *i)* states that lower attention costs lead to a higher level of knowledge reached in equilibrium. According to point *ii)*, interestingly, the overall information may decrease if p_0 increases. Hence, if society considers a measure that delivers public information initially, the incentive effect on the papers that will investigate less as a response may outweigh the first-order effect and lead to less overall information.

The intuition for this last result is as follows. How long the follower investigates is determined by the time the leader can report on an issue with her own information. As the prior precision becomes very small, the leader can attract a lot of attention even with little information gathered previously. The follower can then investigate for a long time.

1.6 Discussion

1.6.1 Experiment-Based vs Belief-Based Modelling

I model information using a signal-/experiment-based approach instead of the commonly used belief-based approach of working directly in the space of distributions over posterior beliefs. See Kamenica and Gentzkow (2011) for static and Ely et al. (2015) for dynamic settings. The experiment-based approach is more convenient for games with multiple senders and dynamic games in particular.

Gentzkow and Kamenica (2016) study a static multiple-sender game with belief-based techniques. They introduce *Blackwell connectedness*, a condition on the information senders can offer. It ensures that each sender can unilaterally deviate to any feasible but more informative posterior distribution. In a simultaneousmove one-shot game, this condition allows them to consider Nash equilibria in which senders choose the same posterior distribution, and no sender has an incentive to deviate to a more informative posterior distribution. In the setting presented here, Blackwell connectedness holds if and only if all senders have one identical signal. Due to the competition, this case is rather uninteresting in my model. There is a competitive equilibrium in which all senders offer all their information in the first round. The receiver chooses randomly which sender to visit, after which all information is observed and the game ends.

Another reason for the experiment-based approach is that different beliefs over signal x_i may arise depending on the information observed previously. Identical offers would give different distributions over posterior beliefs depending on the current belief. Alonso and Camara (2016) identify a bijection between the posteriors that emerge when players with different priors update their beliefs through a common signal. Given that histories are public in my game, this connection would permit to set up the model with the belief-based approach: letting the sender choose the posterior distribution for a baseline belief and keeping track of the resulting posterior distributions for different beliefs. However, with a large set of possible beliefs that can emerge in any round, this is not tractable. Furthermore, the model with experiments can be easily extended to the case with multiple receivers who may have observed different realisations before choosing from the same set of signals, and the case in which senders cannot observe the realisations of their competitors.

1.6.2 Lack of Commitment and Poisson Arrival

I relate my result of the single-sender case to Zhong (2019) and identify the lack of intertemporal commitment by the sender as an additional driver in favour of Poisson processes against Gaussian processes.

Zhong (2019) shows that Poisson learning is uniquely optimal for a decision maker who designs an optimal information process subject to costs proportional to the expected reduction in entropy. His paper shows that this is driven by discounting and the resulting risk preferences. With linear time-costs instead of discounting, Poisson and Gaussian information are both optimal for the decision maker. Featuring no discounting, my model identifies another channel that requires a jump component in the revelation of information – namely, the lack of intertemporal commitment by the sender. Together with the non-concavity in the value of information, this requires discrete jumps in the receiver's belief with positive probability, even in the limit as the length of a period goes to 0.

To illustrate this, consider the following example. The receiver has to guess the outcome of a coin flip, which is the information a single sender holds. That is, $\Omega = X_1 = A = \{0, 1\}$ with $x_1 = \omega$. Let the receiver get utility 1 whenever he guesses correctly and 0 otherwise. Then, for belief $\xi = Pr[\omega = 1]$, we have $U(\xi) = \max\{\xi, 1 - \xi\}$. If the current belief is ξ, Lemma 1 implies that the monopolist gets $(1 - \max\{\xi, 1 - \xi\}) \frac{1}{c}$ rounds of attention. This holds for all rounds and beliefs.[22] For the receiver to be willing to pay the attention cost c in the current round, any experiment has to satisfy

$$\mathbb{E}_m \left[\max\{1 - \xi'(\xi, m), \xi'(\xi, m)\}\right] - c \geq \max\{1 - \xi, \xi\}.$$

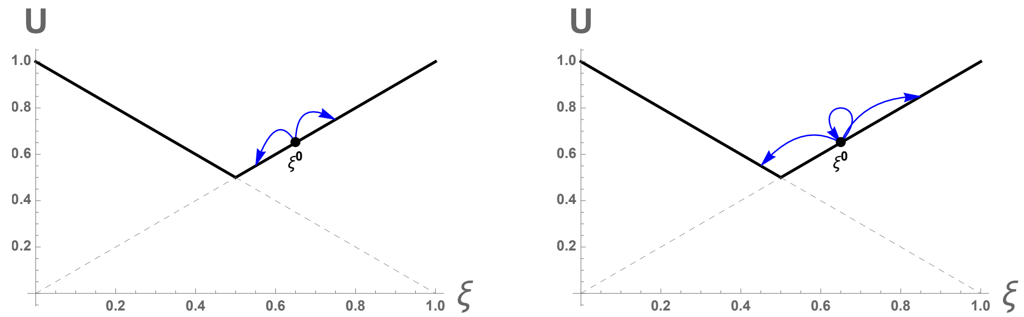
It follows that the message m resulting from the offered experiment has to change the receiver's optimal action with positive probability. If the chosen action stays constant,

$$\begin{aligned}&\mathbb{E}_m \left[\max\{1 - \xi'(\xi, m), \xi'(\xi, m)\}\right] = \max\left\{\mathbb{E}_m \left[1 - \xi'(\xi, m)\right], \mathbb{E}_m \left[\xi'(\xi, m)\right]\right\} \\ &= \max\{1 - \xi, \xi\}\end{aligned}$$

and the experiment's value is 0. Whenever the current belief ξ is different from $\frac{1}{2}$, this implies that, with positive probability, the experiment has to induce a discrete jump in the belief.

[22] As long as $(1 - \max\{\xi, 1 - \xi\}) \geq c$. If this is not fulfilled, no further information transmission is possible.

Figure 1.3: Experiments without and with action change



The left panel in Figure 1.3 shows an experiment with value 0. Since the action is unchanged at both possible posteriors, and U is linear in between, the expected stopping utility is unchanged. To deliver positive value, any experiment has to change the action with positive probability, which implies for the example in the graph that a posterior belief $< \frac{1}{2}$ has to be reached with positive probability, as shown in the right panel. This is certainly true for discrete rounds that require cost $c > 0$. Yet, jumps in the revelation remain necessary, even in the continuous-time limit with attention cost cdt per interval with length dt. Zhong (2019) shows that without loss, any posterior belief process can be decomposed into a Poisson component with jumps and a gradual Gaussian component. Letting the period length go to 0, the probability of a belief change induced by a Gaussian process vanishes exponentially. Together with the above observation, we can conclude that the information offered by the sender has to include at least some jump component, even as periods become arbitrarily small. Note that the reason for Poisson here is different from the risk preferences induced by discounting in Zhong (2019). In the current model without discounting, Poisson is required by lack of intertemporal commitment on the side of the sender.

1.6.3 Complementary Signals and Hold-Up Problem

To illustrate how complementarities in the senders' information hinder transmission in equilibrium, consider two senders. Each sender observes the outcome of an independent, fair coin flip. The receiver has to guess whether the two coins match or not. Let the receiver's utility again be 1 if he guesses correctly and 0 otherwise. In this case, the two signals (coin flips) are perfect complements. In particular, the value of observing one signal without any information about the other is 0.

Suppose that sender 1 has revealed the result of her coin flip. Sender 2 is a monopolist and requires $\frac{1}{c}\frac{1}{2}$ visits in expectation to reveal her information.[23] Anti-

[23] The value of sender 2's information after knowing x_1 is the difference between being able to

cipating this, the receiver's willingness to pay for sender 1's signal is 0. The receiver is not willing to invest a single visit, even if sender 1 offers to reveal her information for sure. The cost $c > 0$ is too high.

There can be no information transmission in equilibrium due to this hold-up problem that arises with one sender after having observed the other sender's information.

Going away from the case where the first sender offers to reveal her information perfectly, suppose that sender 1 revealed partial information, and for concreteness, let the current belief about x_1 be ξ_1 with $\frac{1}{2} < \xi_1 < 1$. Then, the value of sender 2's information is $\bar{v}_2(\xi_1) = \xi_1 - \frac{1}{2}$. After knowing the result of the second coin, the receiver guesses correctly whether they match or not with probability $\xi_1 > \frac{1}{2}$. Sender 2 can extract at least $\bar{v}_2(\xi_1)/c$ units of attention with the corresponding AoN strategy. Note that, as signals are complements, the value of her information will increase in expectation with further revelations about x_1.

1.7 Concluding Remarks

This paper presents a tractable model to study dynamic information provision by senders who are interested in maximising attention. A simple class of processes suffices to transmit all information from senders to the receiver with minimal attention. For the single sender case, I identify the lack of intertemporal commitment as a novel driver for Poisson information. With competition, I identify a condition on the informational externalities that ensures that all information can be transmitted. The concentration of information on fewer senders decreases the receiver-payoff in his preferred equilibrium. In the case of Gaussian information where each sender's informational endowment can be parametrised by a single number, equal precision levels among senders are preferable for the receiver.

If the senders' precision levels are determined in an investigation race, however, they are polarised. The more efficient newspaper exacerbates its informational advantage by investigating longer than the less efficient competitor. An exogenous increase of initially available public information may decrease the newspapers' incentives to investigate enough to decrease the final precision reached in society. Hence, measures that deliver public information on an issue may be counterproductive and lead to less total information on this issue.

The model lends itself to several extensions that are beyond the scope of this paper. While I assumed that all messages are publicly observable, modelling information as experiments can handle heterogeneous priors. Heterogeneous priors may arise if the senders do not observe the message of the visited competitor or if they

guess correctly for sure or with probability $\frac{1}{2}$.

do not observe the receiver's visit history. Incorporating such non-observabilities would allow a welfare comparison between the case in which information providers can track their users across sites and the case in which they are not permitted to do so.

Other interesting avenues for future research are different aspects of information acquisition, such as the choice of issues to report on or the decision between seeking more or less correlation with other newspapers. The tractable computation of equilibrium payoffs in this model can be used as a reduced-form of the payoffs and applied to those questions. Lastly, the introduction of prices in addition to attention allows for comparing membership-based business models to advertisement-based business models.

1.A Appendix

1.A.1 Additional Results and Omitted Proofs

1.A.1.1 Updating of Information

The belief about the state, $\mu(\xi)$, if the signal-belief is ξ, is given by

$$\mu(\xi)(\cdot) = \int_X \mu^0_{|\boldsymbol{x}}(\boldsymbol{x}, \cdot) d\xi(\boldsymbol{x}),$$

where $\mu^0_{|\boldsymbol{x}}$ is the conditional probability of the state ω, given the signals $\boldsymbol{x}$. Note that two different signal-posteriors $\xi \neq \hat{\xi}$ may induce the same state-belief $\mu(\xi)$. As discussed after the proof of Theorem 1, working with μ as the state variable would not, therefore, contain enough information. As the signal space X is complete and separable, a regular conditional probability exists. It has the properties that $\mu^0_{|\boldsymbol{x}}(\cdot, W)$ is measurable for all $W \in \mathcal{B}(\Omega)$ and $\mu^0_{|\boldsymbol{x}}(\boldsymbol{x}, \cdot)$ is a probability measure for all $\boldsymbol{x} \in X$.

At the end of round k, the receiver uses message m_k resulting from the selected experiment λ_k to update the belief from ξ^k to ξ^{k+1}. If, in the following expression, the denominator on the right-hand side is non-zero, the receiver forms ξ^{k+1} through Bayes' rule as follows:

$$\xi^{k+1}(\cdot) = \frac{\int_{\cdot} \lambda_k(x_{d_k}, m_k) d\xi^k(\boldsymbol{x})}{\int_X \lambda_k(x_{d_k}, m_k) d\xi^k(\boldsymbol{x})}.$$

In order to define the updating rule $\xi'(\xi^k, m)$ more generally, note that $L^k(\cdot) \equiv \int_X \lambda_k(x_{d_k}, \cdot) d\xi^k(\boldsymbol{x})$ constitutes a probability measure over M. The updating rule $\xi'(\xi^k, m)$ is the non-negative function that satisfies

$$\int_{M'} \xi'(\xi^k, m)(\cdot) dL^k(m) = \int_{\cdot} \lambda_k(x_{d_k} M') d\xi^k(\boldsymbol{x}), \qquad \text{for all } M' \in \mathcal{B}(M).$$

Such a function ξ' exists and is unique L^k-almost everywhere by the Radon-Nikodym Theorem, as for any $X' \in \mathcal{B}(X)$, the right-hand side, interpreted as a measure on M, is absolutely continuous with respect to L^k (see Billingsley, 1995, p. 422).

An experiment from sender i contains information only about x_i directly, i.e. $\lambda(x_i, \cdot)$ is independent of x_j for all $j \neq i$. However, the receiver's belief about x_j will still change through the correlation among signals. If $d_k = i$, the likelihood ratio for two distinct $\boldsymbol{x}_{-i}, \boldsymbol{x}'_{-i}$, given x_i will not be changed through the updating. That is, $\frac{\xi^{k+1}(x_i, \boldsymbol{x}_{-i})}{\xi^{k+1}(x_i, \boldsymbol{x}'_{-i})} = \frac{\xi^k(x_i, \boldsymbol{x}_{-i})}{\xi^k(x_i, \boldsymbol{x}'_{-i})}$ for all x_i.

Lastly, $v \geq 0$ always, since

$$\begin{aligned} U(\xi) &= \max_{a \in A} \mathbb{E}_{\omega \sim \mu(\xi)} [\, u(a, \omega) \,] \\ &= \max_{a \in A} \mathbb{E}_{m|\xi} \left[\mathbb{E}_{\omega \sim \mu(\xi'(\xi, m))} [\, u(a, \omega) \,] \right] \leq \mathbb{E}_{m|\xi} \left[\max_{a \in A} \mathbb{E}_{\omega \sim \mu(\xi'(\xi, m))} [\, u(a, \omega) \,] \right], \end{aligned}$$

where I shorten the notation from $\mathbb{E}_{x_i \sim \xi} \mathbb{E}_{m \sim \lambda(x_i, \cdot)}$ to $\mathbb{E}_{m|\xi}$. The second equality is due to the martingale property of beliefs, which ensures that $\mathbb{E}_{m|\xi} [\xi'(\xi, m)] = \xi$ for any experiment.

1.A.1.2 Proofs for General Model

Proof of Lemma 1

The AoN equilibrium follows from the text following Lemma 1. Further,

$$\frac{\bar{v}_1\left(\xi^0\right)}{c}$$

is clearly an upper bound for the expected rounds of attention. The AoN strategy ensures the sender this payoff so that she is not willing to deviate to any strategy with a strictly lower payoff. □

Proof of Theorem 1 without Markov Restriction

This proof refers to the proof of Theorem 1, regarding the second step of Claim 1. Let sender i play the AoN strategy from the theorem. By the first step of the claim, the receiver does not stop before sender i's information is revealed. We can therefore determine the number of visits sender i attracts as

$$\mathbb{E}\left[\sum_{n\geq 0}\prod_{m=1}^{n}\left(1-\lambda_i^*(\xi^{k(m)})\right)\right], \tag{1.A.1}$$

where n counts the number of visits to sender i and $k(m)$ is the round in which sender i is visited the m'th time. The process $\left(\mathbb{E}\left[\bar{v}(\xi'(\xi^0,\boldsymbol{x}_{-i}))\;\middle|\;\mathcal{F}_k\right]\right)_{k\geq 0}$ is a martingale. I write the sigma algebra $\mathcal{F}_k$ explicitly instead of the belief ξ^k. The definition (1.1) shows that λ^* is a convex function of the above process, so that the process $\left(\lambda^*(\xi^k)\right)_{k\geq 0}$ is a submartingale. For any finite m, the stopping time $k(m)$ is finite almost surely. Further, $\lambda^*\in[0,1]$, so that the submartingale has bounded increments. This implies that we can apply the optional stopping theorem to derive that, for any $m'>m$:

$$1-\lambda_i^*(\xi^{k(m)})\geq\mathbb{E}\left[\left(1-\lambda_i^*(\xi^{k(m')})\right)\;\middle|\;\mathcal{F}_{k(m)}\right].$$

The following steps show that this permits deriving a lower bound for the number of visits in (1.A.1) given by

$$\sum_{n\geq 0}\left(1-\lambda_i^*(\xi^0)\right)^n=\frac{1}{\lambda_i^*(\xi^0)}. \tag{1.A.2}$$

To see how this is derived, consider, for illustration, the sum in (1.A.1) until $n=2$, which satisfies

$$\begin{aligned}
&\mathbb{E}\left[1+\left(1-\lambda_i^*(\xi^{k(1)})\right)\left(1+\left(1-\lambda_i^*(\xi^{k(2)})\right)\right)\;\middle|\;\mathcal{F}_0\right]\\
&\geq\mathbb{E}\left[1+\mathbb{E}\left[\left(1-\lambda_i^*(\xi^{k(2)})\right)\;\middle|\;\mathcal{F}_{k(1)}\right]\left(1+\left(1-\lambda_i^*(\xi^{k(2)})\right)\right)\;\middle|\;\mathcal{F}_0\right]\\
&=\mathbb{E}\left[1+\mathbb{E}\left[\left(1-\lambda_i^*(\xi^{k(2)})\right)\;\middle|\;\mathcal{F}_{k(1)}\right]\left(1+\mathbb{E}\left[\left(1-\lambda_i^*(\xi^{k(2)})\right)\;\middle|\;\mathcal{F}_{k(1)}\right]\right)\;\middle|\;\mathcal{F}_0\right].
\end{aligned}$$

The inequality uses (1.A.2) and the equality follows from the tower property of conditional expectations. This step can be reiterated. By Doob's martingale convergence theorem, the limit

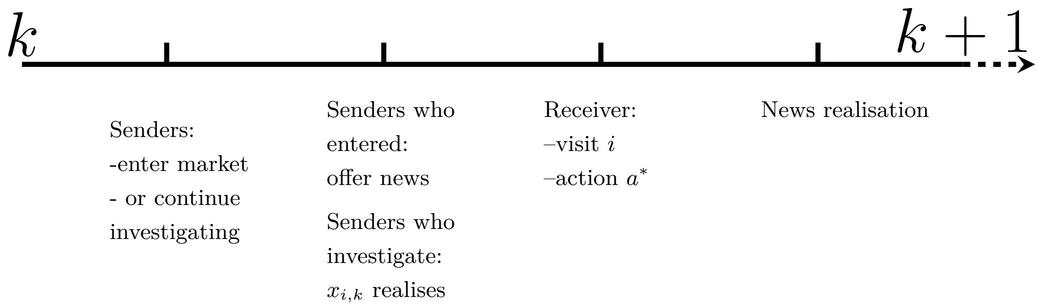
$$\lim_{k\to\infty}\mathbb{E}\left[\left(1-\lambda_i^*(\xi^k)\right)\;\middle|\;\mathcal{F}_{k(m)}\right]$$

exists and is smaller or equal to $\left(1-\lambda_i^*(\xi^0)\right)$. Applying these steps for all $n\in\mathbb{N}_0$, where the submartingale inequality and the tower property have to be used repeatedly for terms with $n>2$, gives the desired result. □

1.A.2 Proofs for Investigation Race

This section presents the discrete time investigation race underlying Section 1.5. As before, the state distribution is $\mathcal{N}(0, 1/p_0)$. The length of each time period is $\Delta > 0$. Each newspaper's precision is determined endogenously in the following stopping game. To obtain information about the state, the papers can investigate before entering the market to disseminate news. Investigating in round n, i.e. from time $n\Delta$ until $(n+1)\Delta$, means that newspaper i is endowed with signal $x_{i,k} \sim \mathcal{N}(\omega, \frac{1}{\rho_i}\frac{1}{\Delta})$. Conditional on the state, signals are independent across senders and rounds. Entering the market allows the newspaper to offer news from that round onward. Note that the normal distribution implies that the signals gathered by sender i from round 0 up to market entry at round n are equivalent to observing a normal signal with precision level $n\rho_i\Delta$. The senders get payoff Δ per round. The receiver's cost is $c\Delta$ and, as mentioned above, he incurs costs only after the first newspaper entered.

Timing



The timing in each round is as follows. Senders decide whether to enter the market. This decision is publicly observed. Senders who entered in this round or before, offer news. For senders who continue investigating, signal $x_{i,k}$ realises. News offers become public, and the receiver decides whether to visit one of the senders who offers news or take the action.

In the main text, I introduced the assumption, $\rho_i > p_0^2 c$, to ensure that each newspaper is efficient enough so that investigation is efficient initially. For period length Δ, the corresponding assumption is that the first round of investigation be efficient:

$$-\frac{1}{p_0 + \rho_i\Delta} - c\Delta > -\frac{1}{p_0} \;\Leftrightarrow\; \rho_i(1 - p_0 c\Delta) > cp_0^2. \tag{1.A.3}$$

Note that this assumption implies that for any length Δ, $p_0 c\Delta < 1$. As Δ goes to zero, (1.A.3) reduces to the assumption in the main text.

As in the main text, I focus on pure strategy equilibria, and I rule out that different entry times are rewarded or punished through the equilibrium that is played after both senders enter the market by focusing on equilibria in which, after both senders are in the market, they play the AoN continuation equilibria corresponding to Theorem 1. However, different from the main text, this section also considers different equilibrium strategies by the leader while she is the only sender in the market.

Results

Lemma A.1. *Suppose the leader has entered the market in round n_ℓ and the follower has not entered. In any equilibrium of the continuation game, the follower does not enter the market before the leader has revealed all her information to the receiver.*

Proof. Playing the AoN equilibria after the follower's market entry implies that her payoff, as a function of entry rounds n_ℓ of the leader and n_f of the follower, is

$$\left(\frac{n_f\rho_f\Delta}{c(p_0+n_\ell\rho_\ell\Delta)(p_0+n_\ell\rho_\ell\Delta+n_f\rho_f\Delta)}\right)\mathbb{1}_{\{\frac{n_f\rho_f}{c(p_0+n_\ell\rho_\ell\Delta)(p_0+n_\ell\rho_\ell\Delta+n_f\rho_f\Delta)}>1\}}.$$

The indicator function whenever

$$\frac{-1}{p_0+n_\ell\rho_\ell\Delta+n_f\rho_f\Delta}-\frac{-1}{p_0+n_\ell\rho_\ell\Delta}>c\Delta.$$

This is condition (A1), ensuring that the information held by the follower is worth at least one round of attention. If the precision of the leader grows too large and her information is revealed too early, the follower cannot attract any attention.

This payoff is increasing in n_f (both the value and the likelihood that the indicator function is one), so that the follower will enter the market only if the receiver would otherwise stop in this round. In the case that the follower does not enter, what makes the receiver stop? If the leader has revealed too much information so that giving out her exact signal is worth less than $c\Delta$, the receiver stops. Waiting one more round and hoping that the follower will enter is not profitable as the follower will extract all the surplus from the receiver. For the leader, it is optimal to replace any such realisations that would lead the receiver to stop (absent entry of the follower) with full information revelation. To see why, consider the leader's payoff in such a round. That is, in a round n in which she offers an experiment λ, such that the set $M^{st}\equiv\{m:\frac{1}{p_0+\rho_\ell n_\ell\Delta(1-\frac{Var(\mathbb{E}[x_\ell|\xi'(\xi^n,m)])}{Var(x_\ell)})}-\frac{1}{p_0+\rho_\ell n_\ell\Delta}<c\Delta\}$ occurs with positive probability. M^{st} includes all messages that make the information about the leader's signal x_ℓ precise enough so that the receiver is not willing to spend a further $c\Delta$, even with the promise of getting all information.[24] Clearly, the leader cannot attract any further visit after a message in M^{st} has realised. The following change in the offered experiment increases the leader's expected utility and ensures that the receiver still accepts. Consider the overall probability of such a message

$$L^n(M^{st})=\int_X\int_{M^{st}}\lambda(x_\ell,dm)\xi^n(d\boldsymbol{x})$$

and replace all messages in this set by revealing no information with probability $\alpha L^n(M^{st})$ and all information with probability $(1-\alpha)L^n(M^{st})$. To ensure the same continuation value for the receiver, α is chosen such that

$$\begin{aligned}&\alpha\frac{-1}{p_0+n_\ell\rho_\ell\Delta(1-\frac{Var(\mathbb{E}[x_\ell|\xi^n])}{Var(x_\ell)})}+(1-\alpha)\frac{-1}{p_0+n_\ell\rho_\ell\Delta}\\&=\frac{1}{L^n(M^{st})}\int_X\int_{M^{st}}\frac{-1}{p_0+\rho_\ell n_\ell\Delta(1-\frac{Var(\mathbb{E}[x_\ell|\xi'(\xi^n,m)])}{Var(x_\ell)})}\lambda(\tilde{x}_\ell,dm)\xi^n(d\tilde{\boldsymbol{x}}).\end{aligned}$$

The right-hand side is the expected value after a message of set M^{st} (note that even with the follower entering, the receiver will be left at her current stopping utility because of the follower's monopoly power). The left-hand side equals this expected utility, either giving no further information or all information held by the leader. Note that $\alpha\in[0,1]$ since for all m, $Var(\mathbb{E}[x_\ell|\xi^n])\leq Var(\mathbb{E}[x_\ell|\xi'(\xi^n,m)])\leq Var(x_\ell)$. The leader is better off, the probability with which this round is her last round of attention decreases, and with positive probability, she reached the next round with the receiver's belief remaining unchanged. □

Knowing that the follower keeps investigating instead of competing actively as long as the leader

[24] Note that the set M^{st} depends on the current belief and the experiment offered.

still holds private information gives rise to the following result, akin to Lemma 1 in the main text with the leader in the role of the monopolist. In every round after she enters, the leader will not offer more information than necessary to keep the receiver indifferent between stopping and visiting.

Lemma A.2. *Suppose the leader enters in round n_ℓ. Let $n_f - 1$ be the round in which all her information is revealed. Then, her expected payoff, $(\mathbb{E}[n_f] - n_\ell)\Delta$, is equal to*

$$(\mathbb{E}[n_f] - n_\ell)\Delta = \frac{1}{c}\left(\frac{1}{p_0} - \frac{1}{p_0 + n_\ell \rho_\ell \Delta}\right) = \frac{n_\ell \rho_\ell \Delta}{c p_0 (p_0 + n_\ell \rho_\ell \Delta)}.$$

This pins down the follower's expected entry time n_f.[25] The realisation of n_f, however, depends on the leader's offer strategy applied from n_ℓ onward. There are several such strategies. The main text focused on the equilibrium in which the leader plays an AoN strategy from n_ℓ onward until her information finally realises. Therefore, I will first consider this equilibrium in what follows and provide the proofs for the main text. Subsequently, I consider different strategies to argue that the effects and results presented in the main text do not hinge on this equilibrium selection.

The AoN offer for the leader entering in round n_ℓ reveals her information in each period with probability

$$\lambda_\ell(n_\ell) = \frac{c p_0 (p_0 + n_\ell \rho_\ell \Delta)}{n_\ell \rho_\ell}.$$

By the assumptions above, this is smaller than one for any $n_\ell \geq 1$.

Fixing the leader's strategy, the payoffs of leader and follower in the investigation race are determined by the identity of the leader ($\ell = 1$ or $\ell = 2$) and the entry time n_ℓ. For the leader, $i = \ell$:

$$L_i(n_\ell) = \frac{1}{c}\frac{n_\ell \rho_i \Delta}{p_0(p_0 + n_\ell \rho_i \Delta)} = \frac{1}{\lambda_i(n_\ell)}. \tag{1.A.4}$$

For the follower, $i = f$ and $j = \ell$:

$$F_i(n_\ell) = \sum_{n_f = n_\ell + 1}^{\infty} (1 - \lambda_j(n_\ell))^{n_f - (n_\ell + 1)} \lambda_j(n_\ell) \frac{1}{c} \frac{n_f \rho_i \Delta \mathbb{1}_{\{\frac{n_f \rho_i}{c(p_0 + n_\ell \rho_j \Delta)(p_0 + n_\ell \rho_j \Delta + n_f \rho_i \Delta)} > 1\}}}{(p_0 + n_\ell \rho_j \Delta)(p_0 + n_\ell \rho_j \Delta + n_f \rho_i \Delta)}. \tag{1.A.5}$$

If both newspapers enter in the same period:

$$\begin{aligned} B_i(n_\ell) =& \mathbb{1}_{\{\frac{1}{c}\frac{n_\ell \rho_i}{(p_0 + n_\ell \rho_j \Delta)(p_0 + n_\ell(\rho_1 + \rho_2)\Delta)} > 1\}} \frac{1}{c} \frac{n_\ell \rho_i \Delta}{(p_0 + n_\ell \rho_j \Delta)(p_0 + n_\ell(\rho_1 + \rho_2)\Delta)} \\ &+ (1 - \mathbb{1}_{\{\frac{1}{c}\frac{n_\ell \rho_i}{(p_0 + n_\ell \rho_j \Delta)(p_0 + n_\ell(\rho_1 + \rho_2)\Delta)} > 1\}}) \max\{0, \frac{1}{c} \frac{n_\ell \rho_i \Delta}{(p_0 + n_\ell \rho_j \Delta)(p_0 + n_\ell(\rho_1 + \rho_2)\Delta)} \\ &+ \Delta - \frac{1}{c} \frac{n_\ell \rho_j \Delta}{(p_0 + n_\ell \rho_i \Delta)(p_0 + n_\ell(\rho_1 + \rho_2)\Delta)}\}. \end{aligned}$$

For the follower i, each round $n_f > n_\ell$ is reached with probability $(1 - \lambda_j(n_\ell))^{n_f - (n_\ell + 1)}$. If the leader's information hits, which happens with probability $\lambda_j(n_\ell)$, the follower receives attention that makes the receiver indifferent between only the leader's or both the leader's and the follower's information. However, as mentioned above, this is only if the difference is worth at least one visit, as captured by the indicator function.

The next results show that if Δ is small enough, there cannot be an equilibrium in which both papers enter the market at the same time, unless both enter in the very first round.

[25] For completeness, I call n_f the follower's entry time, even if the indicator function above is 0 and she cannot attract any attention.

Lemma A.3. *There exist $\epsilon > 0$ such that for any $\Delta \leq \epsilon$, for all $n > 1$, for $i = 1, 2$:*

$$F_i(n) > B_i(n).$$

Proof.

$$\begin{aligned} & F_i(n) > B_i(n) \\ \Longleftarrow & \sum_{n_f=n+1}^{\infty} (1-\lambda_j(n))^{n_f-(n+1)} \lambda_j(n) \frac{n_f \rho_i \Delta}{(p_0 + n\rho_j\Delta)(p_0 + n\rho_j\Delta + n_f\rho_i\Delta)} \\ & \quad > \frac{n\rho_i\Delta}{(p_0 + n\rho_j\Delta)(p_0 + n(\rho_1+\rho_2)\Delta)}. \end{aligned}$$

Note that the second line inequality is sufficient because in the cases where the indicator function in any term in F is 0, it is also 0 for B. In the opposite case, one of the claimed inequalities is always fulfilled. Furthermore, the fraction is increasing in n_f, so that any term of the left-hand sum is greater than the right-hand fraction. It is multiplied with the probability function of a geometric distribution, which sums to one so LHS has to be greater than RHS. □

This shows that there cannot be a pure strategy equilibrium in which the senders enter the market in the same round. The next result shows that if Δ is small enough, the follower's payoff in early periods strictly exceeds the leader's payoff for both players. This, together with the fact that the leader's payoff is increasing in n_ℓ, implies that entering as the leader is strictly dominated by investigating in early periods.

Lemma A.4. *There exist $\epsilon > 0$ such that for any $\Delta \leq \epsilon$, for both papers, $i = 1$ and $i = 2$: there exists n_i^* >1, such that $F_i(n_\ell) > L_i(n_\ell)$ for all $n_\ell < n_i^*$.*

Proof. To take care of the indicator function in F_i, characterised in (1.A.5), note that the term on the LHS of the inequality is increasing in n_f. Therefore, we can define

$$\underline{n}_j(n_\ell) \equiv \min \left\{ n \in \mathbb{N} \;\middle|\; n \geq n_\ell + 1 \;\cap\; \frac{n\rho_j}{c(p_0 + n_\ell\rho_i\Delta)(p_0 + n_\ell\rho_i\Delta + n\rho_j\Delta)} > 1 \right\}.$$

We can write $L_i(n_\ell) < F_i(n_\ell)$ as

$$\frac{n_\ell\rho_i\Delta}{p_0(p_0 + n_\ell\rho_i\Delta)} < \sum_{n_f=\underline{n}_i(n_\ell)}^{\infty} (1-\lambda_j(n_\ell))^{n_f-(n_\ell+1)} \lambda_j(n_\ell) \frac{n_f\rho_i\Delta}{(p_0 + n_\ell\rho_j\Delta)(p_0 + n_\ell\rho_j\Delta + n_f\rho_i\Delta)}.$$

The last fraction in the sum satisfies

$$\frac{n_f\rho_i\Delta}{(p_0 + n_\ell\rho_j\Delta)(p_0 + n_\ell\rho_j\Delta + n_f\rho_i\Delta)} = \frac{1}{p_0 + n_\ell\rho_j\Delta}\left(1 - \frac{p_0 + n_\ell\rho_j\Delta}{p_0 + n_\ell\rho_j\Delta + n_f\rho_i\Delta}\right).$$

The entire sum is then equal to

$$\begin{aligned}
&\sum_{n_f=\underline{n}_i(n_\ell)}^{\infty} (1-\lambda_j(n_\ell))^{n_f-(n_\ell+1)} \lambda_j(n_\ell) \frac{1}{p_0+n_\ell\rho_j\Delta} \\
&- \sum_{n_f=\underline{n}_i(n_\ell)}^{\infty} (1-\lambda_j(n_\ell))^{n_f-(n_\ell+1)} \lambda_j(n_\ell) \frac{1}{p_0+n_\ell\rho_j\Delta+n_f\rho_i\Delta} \\
&= (1-\lambda_j(n_\ell))^{\underline{n}_i(n_\ell)-(n_\ell+1)} \lambda_j(n_\ell) \frac{1}{p_0+n_\ell\rho_j\Delta} \sum_{m=0}^{\infty} (1-\lambda_j(n_\ell))^m \\
&- \sum_{n_f=\underline{n}_i(n_\ell)}^{\infty} (1-\lambda_j(n_\ell))^{n_f-(n_\ell+1)} \lambda_j(n_\ell) \frac{1}{p_0+n_\ell\rho_j\Delta+n_f\rho_i\Delta} \\
&= (1-\lambda_j(n_\ell))^{\underline{n}_i(n_\ell)-(n_\ell+1)} \\
&\left(\frac{1}{p_0+n_\ell\rho_j\Delta} - \lambda_j(n_\ell) \sum_{m=0}^{\infty} (1-\lambda_j(n_\ell))^m \frac{1}{p_0+n_\ell\rho_j\Delta+(n_\ell+1)\rho_i\Delta+m\rho_i\Delta} \right).
\end{aligned}$$

If Δ and n_ℓ are small enough, we get $\underline{n}_i(n_\ell) = n_\ell + 1$ so that the above term is larger than

$$\left(\frac{1}{p_0+n_\ell\rho_j\Delta} - \frac{1}{p_0+n_\ell\rho_j\Delta+(n_\ell+1)\rho_i\Delta} \right) = \frac{(n_\ell+1)\rho_i}{(p_0+n_\ell\rho_j\Delta)(p_0+n_\ell\rho_j\Delta+(n_\ell+1)\rho_i\Delta)}.$$

With this, a sufficient condition for $L_i(n_\ell) < F_i(n_\ell)$ is

$$\frac{n_\ell+1}{n_\ell} \frac{p_0}{p_0+n_\ell\rho_j\Delta} \frac{p_0+n_\ell\rho_i\Delta}{p_0+n_\ell\rho_j\Delta+(n_\ell+1)\rho_i\Delta} > 1,$$

which is satisfied for Δ small enough as the second and third fraction become arbitrarily close to 1 as Δ decreases to 0. For fixed Δ, if $\underline{n}_j(n_i) = n_i + 1$, then $\underline{n}_j(n) = n + 1$ for all $n \leq n_i$. The existence of an n_i^* as in the lemma follows as the term above is decreasing in n_ℓ. □

Considering the limit of (1.A.4) and (1.A.5) as Δ goes to 0 and n_ℓ goes to ∞ fixing the time $k = n\Delta$, we get that

$$L_i(k_\ell) = \frac{1}{\lambda_i(k_\ell)} = \frac{k_\ell\rho_i}{cp_0(p_0+k_\ell\rho_i)},$$

and

$$\begin{aligned}
F_i(k_\ell) &= \frac{1}{p_0+k_\ell\rho_j} - \int_0^{\infty} e^{-k\lambda_j(k_\ell)} \lambda_j(k_\ell) \frac{1}{p_0+k_\ell(\rho_j+\rho_i)+k\rho_i} \\
&= \frac{1}{p_0+k_\ell\rho_j} - \frac{\lambda_j(k_\ell)}{\rho_i} e^{\frac{\lambda_j(k_\ell)}{\rho_i}(p_0+k_\ell(\rho_i+\rho_j))} \int_{\frac{\lambda_j(k_\ell)}{\rho_i}(p_0+k_\ell(\rho_i+\rho_j))}^{\infty} \frac{e^{-s}}{s} ds.
\end{aligned}$$

Proof of Theorem 2

The next results on F and L prove Theorem 2. With the above characterisation, we have for both i that

$$\lim_{k_\ell \downarrow 0} L_i(k_\ell) = \lim_{k_\ell \downarrow 0} F_i(k_\ell) = 0.$$

Taking the derivative with respect to k_ℓ and considering the limit, gives:

$$\lim_{k\downarrow 0}\left(\frac{\partial L_i(k)}{\partial k}\right) = \frac{\rho_i}{cp_0^2} \qquad \text{and}$$
$$\lim_{k\downarrow 0}\left(\frac{\partial F_i(k)}{\partial k}\right) = \frac{\rho_i}{cp_0^2} + \frac{\rho_i\rho_j}{c^2p_0^4}.$$

Hence, the follower's payoff is higher initially. Furthermore,

$$\lim_{k_\ell\to\infty} L_i(k_\ell) = \frac{1}{cp_0} \quad \text{and}$$
$$\lim_{k_\ell\to\infty} F_i(k_\ell) = 0.$$

This shows that F_i crosses L_i from above at least once. To show that this happens at only one $k_\ell > 0$, I show that the derivative of $F_i(k_\ell) - L_i(k_\ell)$ crosses 0 at most twice. This is sufficient to rule out a second positive intersection point since we have established $\lim_{k\ell\downarrow 0} \frac{\partial}{\partial k_\ell}(F_i(k_\ell) - L_i(k\ell)) > 0$ and $\lim_{k\ell\to\infty}(F_i(k_\ell) - L_i(k\ell)) < 0$. At the first intersection, F_i crosses L_i from above. If there were a second intersection point, F_i would again lie above L_i. For $\lim_{k\ell\to\infty}(F_i(k_\ell) - L_i(k\ell)) < 0$ to hold, this would require a third intersection which, in turn, requires that the derivative be 0 at least three times. Define

$$\psi_i(k_\ell) \equiv \frac{cp_0(p_0 + k_\ell\rho_j)(p_0 + k_\ell(\rho_i + \rho_j))}{k_\ell\rho_i\rho_j},$$

and consider

$$\begin{aligned}
&F_i(k_\ell) - L_i(k_\ell)\\
&= \frac{1}{c(p_0 + k_\ell\rho_j)} - \frac{p_0(p_0 + k_\ell\rho_j)}{k_\ell\rho_i\rho_j} e^{\psi_i(k_\ell)} \int_{\psi_i(k_\ell)}^{\infty} \frac{e^{-s}}{s} ds - \frac{k_\ell\rho_i}{cp_0(p_0 + k_\ell\rho_i)}\\
&= \frac{p_0^2 - k_\ell^2\rho_i\rho_j}{cp_0(p_0 + k_\ell\rho_i)(p_0 + k_\ell\rho_j)} - \frac{p_0(p_0 + k_\ell\rho_j)}{k_\ell\rho_i\rho_j} e^{\psi_i(k_\ell)} \int_{\psi_i(k_\ell)}^{\infty} \frac{e^{-s}}{s} ds
\end{aligned}$$

Multiply the term with

$$\frac{k_\ell\rho_i\rho_j}{p_0(p_0 + k_\ell\rho_j)} e^{-\psi_i(k_\ell)} > 0.$$

and consider the derivative

$$\begin{aligned}
&\frac{\partial}{\partial k_\ell}\left(\frac{(p_0^2 - k_\ell^2\rho_i\rho_j)(p_0 + k_\ell(\rho_i + \rho_j))}{p_0(p_0 + k_\ell\rho_i)(p_0 + k_\ell\rho_j)} \frac{e^{-\psi_i(k_\ell)}}{\psi_i(k_\ell)} - \int_{\psi_i(k_\ell)}^{\infty} \frac{e^{-s}}{s} ds\right)\\
&= e^{-\psi_i(k_\ell)} \frac{c(p_0 + k_\ell\rho_i)(p_0 + k_\ell\rho_j)\left(p_0^2 - k_\ell^2\rho_i\rho_j\right)\left(p_0^2 - k_\ell^2\rho_j(\rho_i + \rho_j)\right)}{ck_\ell p_0(p_0 + k_\ell\rho_i)^2(p_0 + k_\ell\rho_j)^3}\\
&\quad - e^{-\psi_i(k_\ell)} \frac{k_\ell\rho_i\rho_j\left(k_\ell^3\rho_i\rho_j(2\rho_i + \rho_j) + 5k_\ell^2 p_0\rho_i\rho_j + k_\ell p_0^2\rho_j - p_0^3\right)}{ck_\ell p_0(p_0 + k_\ell\rho_i)^2(p_0 + k_\ell\rho_j)^3}\\
&\quad + e^{-\psi_i(k_\ell)} \frac{(p_0^2 - k_\ell^2\rho_j(\rho_i + \rho_j)}{k_\ell(p_0 + k_\ell\rho_j)(p_0 + k_\ell(\rho_i + \rho_j))}.
\end{aligned}$$

This is equal to 0 iff

$$\begin{aligned}&\frac{c(p_0+k_\ell\rho_i)(p_0+k_\ell\rho_j)\left(p_0^2-k_\ell^2\rho_i\rho_j\right)\left(p_0^2-k_\ell^2\rho_j(\rho_i+\rho_j)\right)}{cp_0(p_0+k_\ell\rho_i)^2(p_0+k_\ell\rho_j)^2}\\&-\frac{k_\ell\rho_i\rho_j\left(k_\ell^3\rho_i\rho_j(2\rho_i+\rho_j)+5k_\ell^2p_0\rho_i\rho_j+k_\ell p_0^2\rho_j-p_0^3\right)}{cp_0(p_0+k_\ell\rho_i)^2(p_0+k_\ell\rho_j)^2}\\&+\frac{p_0^2-k_\ell^2\rho_j(\rho_i+\rho_j)}{(p_0+k_\ell(\rho_i+\rho_j))}=0.\end{aligned}$$

Multiplying and rearranging yields

$$\begin{aligned}&(p_0+k_\ell(\rho_i+\rho_j))\left(c(p_0+k_\ell\rho_i)(p_0+k_\ell\rho_j)\left(p_0^2-k_\ell^2\rho_i\rho_j\right)\left(p_0^2-k_\ell^2\rho_j(\rho_i+\rho_j)\right)\right)\\&-(p_0+k_\ell(\rho_i+\rho_j))\left(k_\ell\rho_i\rho_j\left(k_\ell^3\rho_i\rho_j(2\rho_i+\rho_j)+5k_\ell^2p_0\rho_i\rho_j+k_\ell p_0^2\rho_j-p_0^3\right)\right)\\&+cp_0(p_0+k_\ell\rho_i)^2(p_0+k_\ell\rho_j)^2\left(p_0^2-k_\ell^2\rho_j(\rho_i+\rho_j)\right)=0.\end{aligned}$$

Collecting k_ℓ with equal exponents gives

$$\begin{aligned}&+ck_\ell^6\rho_i\rho_j^2(\rho_i+\rho_j)^2+k_\ell^5\left(cp_0\rho_j^2(\rho_i+\rho_j)^2+cp_0\rho_i\rho_j(\rho_i+3\rho_j)(\rho_i+\rho_j)\right)\\&+k_\ell^4\left(cp_0^2\rho_i\rho_j(\rho_i+\rho_j)+cp_0^2\rho_j(\rho_i+\rho_j)(\rho_i+3\rho_j)-\rho_i\rho_j(\rho_i+\rho_j)(2\rho_i+\rho_j)\right)\\&+k_\ell^3\left(-cp_0^3\rho_i^2-3cp_0^3\rho_i\rho_j+cp_0^3\rho_j(\rho_i+\rho_j)-p_0\rho_i\rho_j(7\rho_i+6\rho_j)\right)\\&+k_\ell^2\left(-3cp_0^4\rho_i-3cp_0^4\rho_j-p_0^2\rho_j(6\rho_i+\rho_j)\right)+k_\ell\left(p_0^3\rho_i-2cp_0^5\right)+p_0^4=0.\end{aligned}$$

The factor after k_ℓ^3 is negative since $\rho_i > cp_0^2$ for both i. Therefore, there are exactly two 'sign changes' in the sequence, and by the rule of signs, the number of positive roots is at most two. Next, I show that the intersection point k_i^* as defined in Theorem 2 is lower for the less efficient newspaper (with lower ρ_i). For this, consider again the equation $F_i - L_i = 0$. This is equivalent to

$$\frac{(p_0^2-k_\ell^2\rho_i\rho_j)(p_0+k_\ell(\rho_i+\rho_j))}{p_0(p_0+k_\ell\rho_i)(p_0+k_\ell\rho_j)}=\psi_i(k_\ell)e^{\psi_i(k_\ell)}\int_{\psi_i(k_\ell)}^{\infty}\frac{e^{-s}}{s}ds. \tag{1.A.6}$$

with

$$\psi_i(k_\ell)=\frac{cp_0(p_0+k_\ell\rho_j)(p_0+k_\ell(\rho_i+\rho_j))}{k_\ell\rho_i\rho_j}.$$

Note that the left-hand side is symmetric, so that it is the same whether $i = 1$ or $i = 2$. Further, $\psi_i(k_\ell) > \psi_j(k_\ell)$ if and only if $\rho_i < \rho_j$. The function on the right-hand side, $xe^x \int_x^\infty \frac{e^{-s}}{s} ds$, is increasing in x, so that the term on the right crosses the term on the left (from below) at an earlier k_ℓ for higher ψ_i. Hence, the higher efficiency newspaper has a later intersection point. This concludes the proof of Theorem 2. □

Proof of Lemma 3

To establish Lemma 3, suppose wlog that $\rho_1 < \rho_2$, so that paper 1 is the leader and enters the market at $k^* = k_2^*$. To avoid cluttering the notation, I will drop the follower's index from the stopping time and auxiliary function. That is, $k^* = k_2^*$ and $\psi = \psi_2$ for the remainder of the paper. This allows me to write partial derivatives as, for example, ψ_c. Consider the expected precision with which the receiver stops. By Lemma 2, the two papers extract all surplus from the receiver, so that the expected utility from the action, $\mathbb{E}_{k_2}\left[\frac{-1}{p_0+\rho_1 k^*+k_2\rho_2}\right]$, is equal to

$$\frac{-1}{p_0}+c\left(L_1(k^*)+F_2(k^*)\right).$$

Using the formula for L_1 given above and the fact that k^* was defined such that $F_2(k^*) = L_1(k^*)$, this simplifies to

$$\frac{-1}{p_0} + \frac{k^*\rho_1}{p_0(p_0+k^*\rho_1)} + \frac{k^*\rho_2}{p_0(p_0+k^*\rho_2)} = -\frac{p_0^2 - (k^*)^2\rho_1\rho_2}{p_0(p_0+k^*\rho_1)(p_0+k^*\rho_2)}$$

Part i). First, consider the change in precision caused by c:

$$\frac{d}{dc}\left(-\frac{p_0^2 - k^*(c)^2\rho_1\rho_2}{p_0(p_0+k^*(c)\rho_1)(p_0+k^*(c)\rho_2)}\right) = k_c \frac{4kp_0\rho_1\rho_2 + p_0^2(\rho_1+\rho_2) + k^2\rho_1\rho_2(\rho_1+\rho_2)}{(p_0+k\rho_1)^2(p_0+k\rho_2)^2}\Bigg|_{k=k^*}. \tag{1.A.7}$$

Hence, to show that the total precision is decreasing in c, it is sufficient to show that $k_c < 0$.

By the implicit function theorem and the definition of k^* in (1.A.6), we can determine the partial derivative k_c as

$$k_c = \frac{-\frac{\partial}{\partial c}\left(\frac{(p_0^2-k^2\rho_1\rho_2)(p_0+k(\rho_1+\rho_2))}{p_0(p_0+k\rho_1)(p_0+k\rho_2)}\right)\psi + \left(\frac{(p_0^2-k^2\rho_1\rho_2)(p_0+k(\rho_1+\rho_2))}{p_0(p_0+k\rho_1)(p_0+k\rho_2)}\right)\psi_c(\psi+1) - \psi_c\psi}{\frac{\partial}{\partial k}\left(\frac{(p_0^2-k^2\rho_1\rho_2)(p_0+k(\rho_1+\rho_2))}{p_0(p_0+k\rho_1)(p_0+k\rho_2)}\right)\psi - \left(\frac{(p_0^2-k^2\rho_1\rho_2)(p_0+k(\rho_1+\rho_2))}{p_0(p_0+k\rho_1)(p_0+k\rho_2)}\right)\psi_k(\psi+1) + \psi_k\psi}\Bigg|_{k=k^*},$$

Recall that the condition used to determine k^* was $F_2 - L_2 = 0$. As established above, F_2 crosses L_2 from above, so that the denominator of the last expression is negative at $k = k^*$.

It follows that $k_c < 0$ if and only if

$$\begin{aligned} &-\frac{\partial}{\partial c}\left(\frac{(p_0^2-k^2\rho_1\rho_2)(p_0+k(\rho_1+\rho_2))}{p_0(p_0+k\rho_1)(p_0+k\rho_2)}\right)\psi \\ &+\left(\frac{(p_0^2-k^2\rho_1\rho_2)(p_0+k(\rho_1+\rho_2))}{p_0(p_0+k\rho_1)(p_0+k\rho_2)}\right)\psi_c(\psi+1) - \psi_c\psi > 0 \\ \iff &\frac{(p_0^2-k^2\rho_1\rho_2)(p_0+k(\rho_1+\rho_2))}{p_0(p_0+k\rho_1)(p_0+k\rho_2)} > \frac{\psi}{\psi+1} \end{aligned}$$

Applying again the definition of k^*, the fraction on the left is equal to $\psi e^{\psi}\int_{\psi}\frac{e^{-s}}{s}ds$. The exponential integral satisfies the equation $\psi e^{\psi}\int_{\psi}\frac{e^{-s}}{s}ds > \frac{\psi}{\psi+1}$. The comparative static on c follows.

Part ii). To determine the change in p_0, consider the definition of k^*:

$$k^* = \left\{k > 0 \ : \ \frac{(p_0^2-k^2\rho_1\rho_2)(p_0+k(\rho_1+\rho_2))}{p_0(p_0+k\rho_1)(p_0+k\rho_2)} = \psi e^{\psi}\int_{\psi}^{\infty}\frac{e^{-s}}{s}ds\right\}. \tag{1.A.8}$$

The expected final precision is given by

$$-\frac{(p_0^2-k^2\rho_1\rho_2)}{p_0(p_0+k\rho_1)(p_0+k\rho_2)}.$$

Given the definition of k^*, we have

$$-\frac{(p_0^2-k^2\rho_1\rho_2)}{p_0(p_0+k\rho_1)(p_0+k\rho_2)} = -\frac{\psi}{(p_0+k(\rho_1+\rho_2))}e^{\psi}\int_{\psi}^{\infty}\frac{e^{-s}}{s}ds.$$

Note that $\frac{\psi}{(p_0+k(\rho_1+\rho_2))} = \frac{cp_0(p_0+k\rho_1)}{k\rho_1\rho_2} = \frac{cp_0^2}{k\rho_1\rho_2} + \frac{cp_0}{\rho_2}$ goes to zero fast enough so that we must have

$$\lim_{p_0\downarrow 0}\left(-\frac{(p_0^2-k^2\rho_1\rho_2)}{p_0(p_0+k\rho_1)(p_0+k\rho_2)}\right) = 0.$$

Note that the disutility is weakly negative everywhere. For the term to remain at zero, we need $k = \frac{p}{\sqrt{\rho_1\rho_2}}$ everywhere. This, however, gives strictly positive values on the right-hand side of the definition of k^*.

By continuity, this implies that there must be a $\underline{p} > 0$, such that

$$\frac{d}{dp_0}\left(-\frac{p_0^2 - k(p_0)^2\rho_1\rho_2}{p_0(p_0 + k(p_0)\rho_1)(p_0 + k(p_0)\rho_2)}\right) < 0$$

for all $p_0 \le \underline{p}$. □

Gradual Information Release by Leader

While the leader's payoff is fully determined by k_ℓ in any pure strategy equilibrium, the follower's payoff depends on the leader's revelation strategy. The analysis above considered the case where the leader makes AoN offers. However, the leader could choose different distributions, as long as $\mathbb{E}[k_f] - k_\ell = \frac{k_\ell\rho_\ell}{cp_0(p_0+k_\ell\rho_\ell)}$ and the receiver is willing to pay attention in each round. This subsection considers the equilibrium with maximal information precision. As the senders extract all surplus from the receiver and the leader's payoff is fixed for fixed k_ℓ, this is equivalent to maximising the follower's payoff.[26]

Fixing k_ℓ and the resulting $\mathbb{E}[k_f]$, the expected payoff of the follower is

$$\mathbb{E}\left[\frac{k_f\rho_f}{c(p_0 + k_\ell\rho_\ell)(p_0 + k_\ell\rho_\ell + k_f\rho_f)}\right].$$

This is maximised at minimal variance of k_f. With this, we can characterise the information-maximal equilibrium of the stopping game. The payoffs in the stopping game for the leader are

$$L_i(k_\ell) = \frac{1}{c}\frac{k_\ell\rho_i}{p_0(p_0 + k_\ell\rho_i)},$$

and for the follower

$$F_i(k_\ell) = \mathbb{E}\left[\frac{1}{c}\frac{k_2(k_\ell)\rho_j}{(p_0 + k_\ell\rho_i)(p_0 + k_\ell\rho_i + k_2(k_\ell)\rho_j)}\right].$$

In the information-maximal equilibrium, k_2 takes values $k_\ell + \lfloor\frac{k_\ell\rho_1}{cp_0(p_0+k_\ell\rho_1)}\rfloor$ and $k_\ell + \lceil\frac{k_\ell\rho_1}{cp_0(p_0+k_\ell\rho_1)}\rceil$.

In the limit as time periods become small, we get

$$\begin{aligned}F_i(k) &= \frac{1}{c}\frac{(k_\ell + \frac{k_\ell\rho_1}{cp_0(p_0+k_\ell\rho_1)})\rho_j}{(p_0 + k_\ell\rho_i)(p_0 + k_\ell\rho_i + (k_\ell + \frac{k_\ell\rho_1}{cp_0(p_0+k_\ell\rho_1)})\rho_j)} \\ &= \frac{k_\ell\rho_2(cp_0(p_0 + k_\ell\rho_1) + \rho_1)}{c(p_0 + k_\ell\rho_1)(cp_0(p_0 + k_\ell\rho_1)(p_0 + k_\ell(\rho_1 + \rho_2)) + k_\ell\rho_1\rho_2)}.\end{aligned}$$

The leader cannot commit to giving out more information than necessary to attract the receiver's attention. To make sure all her information is revealed by $k_f = k_\ell + \frac{k_\ell\rho_\ell}{cp_0(p_0+k_\ell\rho_\ell)}$ with probability one, she therefore has to release news gradually, so that the receiver is indifferent between stopping and visiting the sender from k_ℓ until k_f. For $k \in [k_\ell, k_f)$, let $\tau(k)$ be the non-decreasing precision level transmitted from the leader to the receiver.The indifference condition prescribes that

[26] Note that maximising precision for a fixed stopping time of the leader does not directly imply that the resulting equilibrium of the investigation race has maximal precision. However, this will be the case here as increasing the follower's payoff across all k_ℓ leads to an increase in the entry equilibrium level, k_ℓ^*.

at all k:

$$\begin{aligned} -\frac{1}{p_0 + \tau[k]\rho_\ell} - c(k - k_\ell) &= -\frac{1}{p_0} \\ \Leftrightarrow \ \tau[k] &= \frac{cp_0^2(k - k_\ell)}{\rho_\ell(1 - cp_0(k - k_\ell))}. \end{aligned}$$

The information that the leader gives out per instant is then $\tau'[k]$, which increases gradually as time passes from k_ℓ to k_f. The more informed the receiver is, the faster his precision has to increase to keep him from stopping.

References

Alonso, R. and Camara, O. (2016). Bayesian Persuasion with Heterogeneous Priors, *Journal of Economic Theory* **165**: 672–706.

Anderson, S. P. and Renault, R. (2009). Comparative Advertising: Disclosing Horizontal Match Information, *RAND Journal of Economics* **40**(3): 558–581.

Angeletos, G.-M. and Pavan, A. (2007). Efficient Use of Information and Social Value of Information, *Econometrica* **75**(4): 1103–1142.

Anton, J. J., Biglaiser, G. and Vettas, N. (2014). Dynamic Price Competition with Capacity Constraints and a Strategic Buyer, *International Economic Review* **55**(3): 943–958.

Au, P. H. (2015). Dynamic Information Disclosure, *RAND Journal of Economics* **46**(4): 791–823.

Ball, I. (2019). Dynamic Information Provision: Rewarding the Past and Guiding the Future. Working Paper.

Bergemann, D. and Bonatti, A. (2019). Markets for Information: An Introduction, *Annual Review of Economics* **11**.

Bergemann, D., Bonatti, A. and Smolin, A. (2018). The Design and Price of Information, *American Economic Review* **108**(1): 1–48.

Bergemann, D. and Morris, S. (2013). Robust Predictions in Games With Incomplete Information, *Econometrica* **81**(4): 1251–1308.

Bergemann, D. and Välimäki, J. (2006). Dynamic Price Competition, *Journal of Economic Theory* **127**(1): 232–263.

Besley, T. and Prat, A. (2006). Handcuffs for the Grabbing Hand? Media Capture and Government Accountability, *American Economic Review* **96**(3): 720–736.

Billingsley, P. (1995). *Probability and Measure*, Wiley Series in Probability and Statistics. Wiley, 3rd ed.

Board, S. and Lu, J. (2018). Competitive Information Disclosure in Search Markets, *Journal of Political Economy* **126**(5): 1965–2010.

Börgers, T., Hernando-Veciana, A. and Krähmer, D. (2013). When are signals complements or substitutes?, *Journal of Economic Theory* **148**(1): 165–195.

Che, Y.-K. and Hörner, J. (2017). Recommender Systems as Mechanisms for Social Learning, *Quarterly Journal of Economics* **133**(2): 871–925.

Che, Y.-K., Kim, K. and Mierendorff, K. (2020). Keeping the Listener Engaged: A Dynamic Model of Bayesian Persuasion. Working Paper, arXiv:2003.07338.

Che, Y.-K. and Mierendorff, K. (2019). Optimal Dynamic Allocation of Attention, *American Economic Review* **109**(8): 2993–3029.

Chen, H. and Suen, W. (2019). Competition for Attention in the News Media Market. Working Paper.

Choi, M., Dai, A. Y. and Kim, K. (2018). Consumer Search and Price Competition, *Econometrica* **86**(4): 1257–1281.

Dudey, M. (1992). Dynamic Edgeworth-Bertrand Competition, *The Quarterly Journal of Economics* **107**(4): 1461–1477.

Ely, J. C. (2017). Beeps, *American Economic Review* **107**(1): 31–53.

Ely, J. C. and Szydlowski, M. (2020). Moving the Goalposts, *Journal of Political Economy* **128**(2): 468–506.

Ely, J., Frankel, A. and Kamenica, E. (2015). Suspense and Surprise, *Journal of Political Economy* **123**(1): 215–260.

Fudenberg, D., Strack, P. and Strzalecki, T. (2018). Speed, Accuracy, and the Optimal Timing of Choices, *American Economic Review* **108**(12): 3651–84.

Fudenberg, D. and Tirole, J. (1991). Perfect Bayesian equilibrium and Sequential Equilibrium, *journal of Economic Theory* **53**(2): 236–260.

Galperti, S. and Trevino, I. (2018). Shared Knowledge and Competition for Attention in Information Markets. Working Paper.

Gentzkow, M. and Kamenica, E. (2016). Competition in Persuasion, *Review of Economic Studies* **84**(1): 300–322.

Gentzkow, M. and Shapiro, J. M. (2006). Media Bias and Reputation, *Journal of political Economy* **114**(2): 280–316.

Gossner, O., Steiner, J. and Stewart, C. (2019). Attention please! Working Paper, University of Zurich, Department of Economics.

Guo, Y. and Shmaya, E. (2019). The Interval Structure of Optimal Disclosure, *Econometrica* **87**(2): 653–675.

Kamenica, E. and Gentzkow, M. (2011). Bayesian Persuasion, *American Economic Review* **101**(6): 2590–2615.

Ke, T. T. and Lin, S. (2020). Informational Complementarity, *Management Science* **66**(8): 3699–3716.

Keller, G., Rady, S. and Cripps, M. (2005). Strategic Experimentation with Exponential Bandits, *Econometrica* **73**(1): 39–68.

Lewis, J. and Cushion, S. (2009). The Thirst to be first, *Journalism Practice* **3**(3): 304–318.

Liang, A. and Mu, X. (2020). Complementary Information and Learning Traps, *The Quarterly Journal of Economics* **135**(1): 389–448.

Liang, A., Mu, X. and Syrgkanis, V. (2019). Dynamically Aggregating Diverse Information. Working Paper.

Margaria, C. and Smolin, A. (2018). Dynamic communication with biased senders, *Games and Economic Behavior* **110**: 330–339.

Martínez-de-Albéniz, V. and Talluri, K. (2011). Dynamic Price Competition with Fixed Capacities, *Management Science* **57**(6): 1078–1093.

Mayskaya, T. (2017). Dynamic Choice of Information Sources. Working Paper, California Institute of Technology.

Morris, S. and Shin, H. S. (2002). Social Value of Public Information, *American Economic Review* **92**(5): 1521–1534.

Morris, S. and Strack, P. (2017). The Wald problem and the Equivalence of Sequential Sampling and static Information Costs. Working Paper.

Moscarini, G. and Smith, L. (2001). The Optimal Level of Experimentation, *Econometrica* **69**(6): 1629–1644.

Mullainathan, S. and Shleifer, A. (2005). The Market for News, *American Economic Review* **95**(4): 1031–1053.

Orlov, D., Skrzypacz, A. and Zryumov, P. (2020). Persuading the Principal to Wait, *Journal of Political Economy* **128**(7): 000–000.

Pant, A. and Trombetta, F. (2019). The Newsroom Dilemma. Working Paper, SSRN 3447908.

Perego, J. and Yuksel, S. (2018). Media Competition and Social Disagreement. Working Paper.

Renault, J., Solan, E. and Vieille, N. (2017). Optimal Dynamic Information Provision, *Games and Economic Behavior* **104**: 329–349.

Smolin, A. (2017). Dynamic Evaluation Design. Working Paper.

Sun, M. (2011). Disclosing Multiple Product Attributes, *Journal of Economics & Management Strategy* **20**(1): 195–224.

Vives, X. (1996). Social learning and rational expectations, *European Economic Review* **40**(3-5): 589–601.

Wald, A. (1947). Foundations of a General Theory of Sequential Decision Functions, *Econometrica* **15**(4): 279.

Wolinsky, A. (1986). True Monopolistic Competition as a Result of Imperfect Information, *Quarterly Journal of Economics* **101**(3): 493–511.

Zhong, W. (2019). Optimal Dynamic Information Acquisition. Working Paper.

2

Costless Information and Costly Verification: A Case for Transparency

Joint with Deniz Kattwinkel

2.1 Introduction

A principal has to take a binary decision for which she relies on an agent's private information. The agent prefers one of the two actions independent of his information. Prior to the decision, the principal privately observes a signal about the agent's information. She cannot incentivise the agent through monetary transfers but has the opportunity to reveal his information at a cost.

Examples for this setting include: a human resource department decides whether to hire a candidate, a judge decides whether to acquit or convict a defendant, or a competition authority decides whether to grant or deny a company permission to merge with or acquire another firm.

While one party – the agent – has a clear preference toward one action (the candidate wants to be hired, the defendant wants to be acquitted, and the company wants to merge), the preferences of the other party – the principal – depend on information that is privately held by the agent. Here, one may think of the candidate's ability, the defendant's guilt, or the company's competitive position in the market.

Often, monetary transfers to elicit the agent's private information are not feasible,[1] but the principal can learn the information at a cost, for example, by conducting an assessment centre, a trial, or a market analysis. However, verification is

[1] The assumption is that payments cannot depend on the agent's report. Even though a public sector job entails payments, if the payment is fixed, it cannot be used to incentivise truthful reports of the candidate's ability.

costly, so the principal has an incentive to economise on it.

Typically, costly information acquisition is not the only way to learn the agent's private information. The potential employer receives references or recommendation letters from previous supervisors, the judge can inspect the outcome of pretrial investigations, and the competition authority has sector-specific knowledge derived from its supervisory function. The principal privately observes factors that are correlated with the agent's type.

In this paper, we introduce correlated information to a mechanism design setting with costly state verification. We want to explore how the principal can use her private information to minimise the verification costs and the decision inefficiencies required to elicit the agent's type. Can the principal exploit the fact that her correlated information is secret?[2]

Results. We show that the optimal Bayesian incentive compatible (BIC) mechanism takes a simple *cutoff with appeal* structure: if the principal observes a signal that makes her sufficiently certain that the agent's preferred action is also what she prefers, she takes it, independent of the agent's type report. If the signal falls below this cutoff, she takes the nonpreferred action by default but gives the agent the possibility to appeal. An appeal is always verified and induces the agent-preferred action whenever his type exceeds a threshold. This appeal threshold is set such that the types who appeal are exactly those which make it worthwhile for the principal to implement the agent-preferred action and pay the verification cost.

A remarkable feature of this mechanism is that it is ex-post incentive compatible (EPIC): truthfully reporting his type remains a best response of the agent even if the principal's signal was known to him. Hence, the principal does not benefit from the secrecy of her signal. This clashes with the observation that institutions invest significant cost and effort to keep information private. A potential justification for this investment is the possibility to increase efficiency through private information. When the signal realisation is kept private from the agent, his beliefs about the signal vary with his type due to the correlation. We identify two channels through which the principal could potentially exploit secrecy to increase efficiency by saving verification costs. However, as the main contribution, we show that the *optimal* mechanism (decision *and* verification rule) does not make use of secrecy. Its decision-rule satisfies monotonicity properties that render the two channels futile.

The first potential channel to benefit from secrecy is the redistribution of excess allocation utility across different signal realisations to satisfy Bayesian Incentive Constraints. We show that, as in the case without correlation, pointwise mono-

[2] In contrast to our result outlined below, the literature on mechanisms with transfers suggests that secrecy can be exploited. See Myerson (1981), Crémer and McLean (1988), McAfee and Reny (1992). A detailed discussion of the related literature is contained in Section 2.6.1.

tonicity of the optimal decision-rule in the agent's type rules out this benefit (see Manelli and Vincent, 2010, Gershkov et al., 2013). The second potential channel to benefit from secrecy exploits the belief heterogeneity induced by correlation to fine-tune verification probabilities, so that, for certain types, verification is subjectively more likely (see Crémer and McLean, 1988, for the fine-tuned lotteries in the case with money). We show that pointwise monotonicity in the principal's signal and non-randomness of the optimal decision-rule rule out this benefit.

Our result advocates transparent procedures as the principal does not lose from revealing her private information prior to interacting with the agent. Accordingly, efficiency concerns do not justify withholding information. The human resource department showing the references to the candidate, the judge informing the defendant of pretrial investigation results, or the competition authority publicising her market assessments would not constrain the implementation of the optimal procedures. The cost and effort to keep information private must, therefore, be explained by other motives.[3] Our findings are in line with the development of codes of criminal procedure in continental Europe. While modern codes prescribe the disclosure of all potential charges to the defendant, this was not always the case (see Section 2.3.1 for a discussion of the evolution of the Austrian code of procedure).

After establishing that full transparency is no worse than withholding all information, we argue how this implies that full transparency also weakly dominates any strategic information release or design by the principal.

The equivalence between EPIC and BIC implementability is shown in Manelli and Vincent (2010) and Gershkov, Goeree, Kushnir, Moldovanu and Shi (2013) under the standard mechanism design assumptions. These seminal results depend crucially on the independence of information. With correlated information, this equivalence does not hold generally. Nevertheless, in our setting, the optimal BIC mechanism can be implemented EPIC, despite correlation. To the best of our knowledge, this is the first result of this kind under correlated information.[4] Without the assumption of independent information, standard techniques to characterise implementable mechanisms do not apply. Therefore, we have to use new techniques to characterise the optimal mechanism: we adopt a variational approach.

We extend the analysis to settings in which the principal's signal has a direct effect on her payoffs. When this effect is positive, our result carries over: transparency comes without loss for the principal. If, in contrast, the direct effect is negative, the principal benefits from secrecy. This reveals that correlated information poses a limitation to the general equivalence between BIC and EPIC implementation.

[3] One motive to keep references secret from the candidate are the incentives of the issuer.

[4] While the allocation mechanism in Crémer and McLean (1988) satisfies dominant incentive compatibility in the second stage, the surplus extraction in the first stage requires that agents do not know the realisation of the correlated information.

2.2 Model

In the remainder of the paper, we use the terminology of one specific binary decision: the allocation of a single indivisible good.

2.2.1 Setup

The principal (she) decides whether to allocate the good to the agent (he). Her allocation preferences depend on the agent's private *type* $t \in \mathbb{R}$. The set of possible types, T, is finite.

While t is unknown to the principal, she receives costless information about it in form of a private *signal* $s \in S$, finite and ordered. Type t and signal s are jointly distributed according to $f(t, s) > 0$ for all $t \in T, s \in S$. The signal satisfies the Monotone Likelihood Ratio Property (MLRP): for all $t < t' \in T$, $\frac{f(t',s)}{f(t,s)}$ is nondecreasing in s. This implies that a higher signal is more indicative of a higher type.[5] On top of the costless information, the principal has the option to learn t at *verification* cost $c > 0$. Verification is perfect; she learns the exact type.[6]

The principal derives *valuation* $v(t)$ when allocating the good to an agent of type t. We normalise the value she derives from not allocating to 0. Therefore, v represents the *net* value for the principal. Valuation $v(t)$ is nondecreasing and there are $t', t'' \in T$ with $v(t') < 0 < v(t'')$.[7] When the agent has type t he receives *utility* $u(t) > 0$ from the good. His payoff from not receiving the good is always zero.

2.2.2 Mechanisms

We study the interaction between principal and agent in a mechanism design setting and characterise the mechanism that maximises the principal's expected payoff net of verification costs. The principal can design arbitrary mechanisms and the agent plays a Bayesian best response after learning his type. A key question in this setting is: can the principal use her private information – the signal – to elicit the agent's information – the type? To render the use of her information possible, we assume that the signal is contractible. She can commit to mechanisms that are contingent on the signal realisation.[8] In the appendix, we therefore define a broad class of

[5] MLRP is equivalent to requiring that t and s be affiliated (Milgrom and Weber, 1982, p. 1098)

[6] Whether the verification technology reveals the true type of the agent or just confirms whether the agent has a specific type or not, does not alter our results.

[7] Otherwise, the principal could implement her first-best allocation without the agent's private information.

[8] This approach gives the principal maximal flexibility to use her information. She can commit to truthfully communicating her signal realisation to the mechanism. An alternative approach would be to consider the *informed principal* problem, requiring the mechanism to make truthful communication incentive compatible for her. We show in Section 2.6 that our results are robust to this modelling choice. The optimal mechanism constitutes an equilibrium in the informed principal problem.

dynamic mechanisms that allow the principal to release garblings of her information at any point of their interaction. This covers any potential for information design by the principal. It turns out that this can be captured entirely within a simple class of mechanisms:

A *direct* mechanism in this setting specifies for any type-signal pair (t, s) two probabilities $x(t, s)$ and $z(t, s)$ and proceeds as follows. It asks the agent to report his type. Based on this report t and the signal realisation s one of three distinct events occurs:

1. With probability $x(t, s)$ the good is allocated to the agent and he is not verified.
2. With probability $z(t, s)$ the agent is verified. Then, the good is allocated to him if and only if he is found to have reported truthfully.
3. With probability $1 - x(t, s) - z(t, s)$ the good is not allocated to the agent and he is not verified.

Feasibility requires the *total allocation* probability $x(t, s) + z(t, s)$ not to exceed 1. In the remainder we refer to $x(t, s)$ as the *non-verified allocation* probability. A mechanism is called *truthful* if reporting truthfully is a best response for all types. The theorem below combines a revelation principle with optimality considerations.

Theorem 1. *There is a direct truthful mechanism which maximises the principal's expected valuation net of verification costs.*

In the proof in Appendix 2.A, we first derive a revelation principle (reminiscent of Ben-Porath, Dekel and Lipman (2014) and Akbarpour and Li (2020)) for our setting with correlated information. Then we exploit that any optimal mechanism needs to satisfy two intuitive properties. *Maximal Punishment:* if verification reveals a misreport, the agent does not receive the good. *Minimal Verification*: after his report is verified to be true, the agent receives the good for sure.

Truthful direct mechanisms do not restrict the principal's ability to strategically release information. Take a mechanism that is not in direct form. Suppose the mechanism reveals a garbling of the signal and then asks the agent to send a message. Different realisations of the garbling induce different beliefs about the signal when the agent sends his message. The revelation principle shows that this information-design mechanism can be replicated by a direct mechanism. This direct mechanism asks the agent for his type and then internally simulates the original mechanism with the agent's best response corresponding to the reported type. Although the agent's belief in the direct mechanism is not updated before his report, when evaluating the expected utility from his report, he takes the perspective of the simulated agent. This way, the information design in the original mechanism affects the incentives of the agent to report truthfully in the direct mechanism.

2.2.3 Incentive Compatibility

In standard mechanism design problems, the set of feasible allocations is pinned down by the incentive compatibility constraints through the well-known integral characterisation (Myerson, 1981). Our design setting is non-standard in two ways: the absence of transfers and the presence of correlated information. Ben-Porath, Dekel and Lipman (2014) show how the former impedes the integral characterisation and present a tractable characterisation for mechanisms without transfers and with independent information. Correlated information impedes this approach (see Example 1) and requires an alternative methodology.

Bayesian Incentive Compatibility. Absent monetary transfers, the agent cares solely about the probability of receiving the good. Consider the incentives of an agent of type t. He does not know the signal realisation. If he reports truthfully, he faces the random (as a function of the random variable s) allocation probability $x(t,s)+z(t,s)$. Whether his report is verified is irrelevant for him. If, however, t reports $\hat{t} \neq t$, he receives the good with random probability $x(\hat{t},s)$, i.e. only if he is not verified. Therefore, type t prefers reporting t to reporting $\hat{t}$ if

$$u(t) \cdot \mathbb{E}_s \left[x(t,s) + z(t,s) \mid t \right] \quad \geq \quad u(t) \cdot \mathbb{E}_s \left[x(\hat{t},s) \mid t \right].$$

Since every type derives strictly positive utility from the good ($u(t) > 0$), type t's preference intensity can be eliminated from the IC constraint. The agent simply maximises his expected allocation probability and the Bayesian incentive constraint can be expressed as

$$\mathbb{E}_s \left[\left(x(t,s) + z(t,s) - x(\hat{t},s) \right) \mathbb{1}_{\{t\}} \right] \;=\; \sum_{s \in S} f(t,s) \left[x(t,s) + z(t,s) - x(\hat{t},s) \right] \;\geq 0. \tag{$BIC_{t,\hat{t}}$}$$

In mechanism design without transfers and independent information, normalising the preference intensities implies that the expected utility of any misreport $\hat{t}$ is independent of the agent's true type, t.[9] Therefore, incentive compatibility holds whenever the type with the lowest allocation probability does not want to misreport (see Ben-Porath, Dekel and Lipman, 2014). In our model with correlated information, this does not hold. Different types hold different beliefs about s, so that interim expected allocation probabilities for a given report are not equal for different types. Hence, with correlation, one must consider the interim expectations for all reports for all types. The following example illustrates this and will later be revisited to illustrate the intuition for our results.

[9] Without correlation, the expectation over s is constant in t.

Example 1. *The agent's type is either high or low, $t \in \{L, H\}$. The principal observes signal $s \in \{\ell, h\}$. Type and signal are jointly distributed according to*

$$\begin{pmatrix} f(L,\ell) & f(L,h) \\ f(H,\ell) & f(H,h) \end{pmatrix} = \frac{1}{6}\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}. \tag{2.1}$$

Consider the following non-verified allocation probabilities

$$(x(L,\ell), x(L,h), x(H,\ell), x(H,h)) = (0,1,1,0)\,, \tag{2.2}$$

which allocate if and only if report and signal do not match. Suppose there is no verification. This mechanism is not incentive compatible. Both types prefer to misreport because they put higher probability on the matching signal. For both types, the interim expected allocation probability from truthtelling is $\frac{1}{3}$ while misreporting yields $\frac{2}{3}$.

A general characterisation of incentive compatibility with belief heterogeneity remains an open question. Instead of characterising incentive compatible mechanisms explicitly, Crémer and McLean (1988) and McAfee and Reny (1992) show that, with money, the belief heterogeneity allows the principal to extract all surplus from the agents.[10] Full surplus extraction will not be feasible in our setting. In contrast, we consider the following, more restrictive incentive constraints as an intermediate step toward characterising the optimal mechanism.

Ex-Post Incentive Compatibility and Transparency. We call a mechanism *transparently* implementable if there is an implementation that starts with the principal making her information public. In the case of direct mechanisms, transparency requires that all types t do not misreport after observing any signal s. This is, thus, equivalent to requiring ex-post incentive compatibility (EPIC):

$$x(t,s) + z(t,s) - x(\hat{t},s) \geq 0. \tag{$EPIC(s)_{t,\hat{t}}$}$$

Every Bayesian incentive constraint ($BIC_{t,\hat{t}}$) is a weighted sum of the corresponding ($EPIC(s)_{t,\hat{t}}$) constraints. Incentive compatible transparent mechanisms are therefore BIC. In Example 1, after learning the signal, both types agree which report is most profitable (the report contrary to the signal). With transparency, when all types learn the signal, the belief heterogeneity is resolved. This facilitates the characterisation of optimal transparent mechanisms (Section 2.4.2). Paired with our main conceptual contribution (Section 2.4.1) – that any BIC mechanism can be

[10] They establish the existence of an incentive compatible mechanism that allocates efficiently and extracts all surplus and, therefore, must be optimal.

made transparent without loss for the principal – this yields optimal mechanisms in the larger class of BIC rules.

2.3 Optimal Mechanisms

The principal designs a mechanism that maximises her expected utility from the allocation net of the cost of verification. If the good is assigned without verification, she gains $v(t)$. In the case of allocation with prior verification, she additionally pays cost c. Hence, the principal's problem can be stated as the following linear program:

$$(LP): \quad \max_{(\boldsymbol{x},\boldsymbol{z})\geq 0} \mathbb{E}\left[\, x(t,s)\; v(t) + z(t,s)\; (v(t)-c)\,\right]$$

$$\text{s.t. for all } t, \hat{t} \in T, s \in S: \qquad (BIC_{t,\hat{t}})$$

$$\text{and} \qquad x(t,s) + z(t,s) \leq 1.$$

Note that the principal optimises subject to the Bayesian incentive constraints (not to the stronger transparency constraints). The principal's value from a transparent mechanism cannot exceed the value from the above problem. The following class of mechanisms plays an important role in the ensuing analysis:

Definition 1. *A mechanism $(\boldsymbol{x}, \boldsymbol{z})$ is called* ***cutoff with appeal*** *if there exists a cutoff $\bar{s}$ and an appeal threshold $\bar{t}$ such that:*

i) If $s \geq \bar{s}$, then $x(t,s) = 1$ for all t and $z(t,s) = 0$ for all t.

ii) If $s < \bar{s}$, then $x(t,s) = 0$ for all t and $z(t,s) = \begin{cases} 1 & \text{for } t \geq \bar{t} \\ 0 & \text{for } t < \bar{t}. \end{cases}$

Figure 2.1: Cutoff with appeal

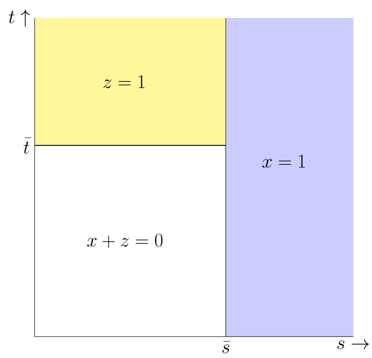


Figure 2.1 sketches a cutoff with appeal mechanism. If the signal realisation is above the cutoff $\bar{s}$, the principal allocates the good to the agent irrespective of his reported type without verification ($x = 1$). If the signal is below the cutoff, the agent

can receive the good only after his type report is verified to be above the threshold $\bar{t}$ ($z = 1$). Note that appeal threshold $\bar{t}$ is the same after all signals. We call this class *cutoff with appeal* because it can be implemented by the following procedure. For signals above $\bar{s}$, the principal allocates without eliciting any information from the agent. For signals below $\bar{s}$, the default is not to allocate but the principal gives the agent the opportunity to appeal. An appeal is granted only after the type is verified to be above the threshold. Our main result shows that the optimal mechanism can be found in this class and specifies the optimal cutoff for the signal and the appeal threshold for the type.

Theorem 2. *The principal's problem is solved by the cutoff with appeal mechanism with cutoff*[11]

$$\bar{s} = \min\left\{ s \mid \mathbb{E}_t[\, v(t) \mid s \,] > \mathbb{E}_t[\, (v(t) - c\,)^+ \mid s \,] \right\},$$

and appeal threshold

$$\bar{t} = \min\left\{ t \mid v(t) - c > 0 \right\}.$$

By the positive correlation (MLRP), higher signals make the principal more optimistic about the agent's type. If the signal exceeds the cutoff $\bar{s}$, she is sufficiently optimistic and allocates without eliciting any further information from the agent. If the signal is below the cutoff, the principal is pessimistic and makes her decision type-dependent. To prevent misreports she allocates only after type verification. The threshold $\bar{t}$ is such that for all higher types the principal profits from allocating even accounting for the verification costs (i.e. when $v(t) - c > 0$). Therefore, given a signal s below the cutoff, the principal's expected value is $\mathbb{E}_t[(v(t) - c)^+|s]$. The optimal cutoff $\bar{s}$ is set such that the principal prefers this value when the signal falls below $\bar{s}$ and prefers the expected value from allocating to all types otherwise. In this optimal mechanism, the principal does not exploit the heterogeneous beliefs for information elicitation. In particular, the agent would report her type truthfully given any belief about the signal. Our cutoff with appeal procedure does not entail complex surplus-extracting schemes as the literature suggests for optimal mechanisms in settings with money (Crémer and McLean, 1988). This does not mean that benefiting from belief heterogeneity is generally impossible in our setting. Section 2.4.1 illustrates how the principal optimally uses type-dependent beliefs in the same setting when her objective is not solely efficiency.

The proof of Theorem 2 is presented in Section 2.4 and consists of two steps. The first step contains the main conceptual contribution of the paper: the prin-

[11] For $r \in \mathbb{R}$ we denote the positive part by $(r)^+ = \max\{r, 0\}$

cipal can achieve her optimum in the class of BIC mechanisms with a transparently implementable mechanism. Thus, transparency entails no loss. The second step completes the proof by showing that the cutoff with appeal mechanism in Theorem 2 is optimal in the class of transparent mechanisms. Before presenting the proof, we collect important features of the optimal mechanism and demonstrate how our result applies to the optimal design of court procedures.

Transparent Implementation. The cutoff with appeal mechanism in Theorem 2 can be implemented transparently. The principal could first reveal her signal to the agent and then ask him to report his type. To see why, consider Figure 2.1. If the signal exceeds the cutoff $\bar{s}$, the allocation is independent of the report (the shaded blue area). If the signal falls below $\bar{s}$, the agent can only get the good after being verified (the shaded yellow area) so that misreporting cannot be beneficial even when the agent knows the signal.

Minimal Communication. The optimal mechanism can be implemented with a minimum of communication between the agent and the principal: the principal first takes a provisional decision on the allocation which is only based on her signal. Then the agent is given the opportunity to appeal against this decision. This is the only instance where the two have to exchange binary messages. In case of appeal, the principal verifies and allocates according to the outcome.

Unique Implementation. If his type is below the threshold, the agent's chances of getting the good are unaffected by his report. Hence, he is indifferent between truthtelling and any misreport. To make truthtelling the unique best response, consider the following amendment of the optimal mechanism. The principal offers a small probability of allocating the good when the agent reports a type below the threshold and the signal falls below the cutoff. Then, the agent has strict incentives to report truthfully for any type. If the probability is chosen small enough this mechanism with strict incentives achieves a payoff arbitrarily close to the optimum.

Futility of Information Design. The contractability of the signal gives the principal maximal flexibility to use her private information. Nevertheless, by our main finding, the optimal mechanism is transparent. This implies that the principal does not profit from persuading the agent to reveal his information by any form of information design. As outlined in the discussion following Theorem 1, our class of direct mechanisms covers procedures in which the principal commits to release parts of her information to the agent to manipulate his beliefs.

With less flexibility, e.g. when the private signal is not contractible, information design remains futile. As a consequence, our mechanism also solves the *informed principal problem* as discussed in Section 2.6.2.

2.3.1 Informing the Defendant

Consider the following application of our model: a judge has to decide whether to acquit or convict a defendant. The defendant privately knows whether he is guilty or innocent. The judge observes the result of a pretrial investigation and can conduct a full trial which is costly but will reveal whether the defendant is guilty. She wants to acquit the defendant if and only if he is innocent whilst the defendant prefers being acquitted irrespective of his guilt. When we model the defendant as the agent, the judge as the principal and decision to acquit as the allocation ($x = 1$) we can identify the agent being innocent with $t = 1$ and $t = 0$ with the agent being guilty. We can capture the preferences of the judge (to only acquit an innocent agent) by $v(t) = t - 1/2$.[12]

The optimal mechanism we derive resembles the proceedings of a pretrial. The case is dismissed if the signal for the defendant's innocence is strong enough, i.e. the charge is weak. If the signal for innocence is below this cutoff, the agent can plead guilty and is convicted, or can request a trial by pleading not guilty, after which he is acquitted if indeed found to be not guilty and convicted otherwise.

A relevant implication of our transparency result is that the justice system does not profit from keeping the discovery of pretrial investigations secret. This is established practice in modern codes of procedures (see Brady v. Maryland, 1963, for the case of U.S. federal law)[13] but was not always the case. Compare for example today's Austrian criminal code of procedure (StPO, §6 (2)) with the code of 1803 (Franz II, 1803, §331). While the modern code grants the right to learn about all potential charges to the defendant, the version from 1803 grants much more discretion in the extent of information released to the defendant, stating that he has to be informed only as far as necessary to notify him that he is accused.[14]

If one extends the model and allows for more types of the agent, i.e. guilty ($t = 0$), guilty of a minor crime ($t = 3/4$) and, innocent ($t = 1$) another feature of our optimal mechanism arises in the pretrial application. When a defendant who is

[12] If an innocent agent ($t = 0$) is convicted ($x = 0$) the net utility loss is given by $0 - v(1) = -1/2$. If a guilty agent ($t = 0$) is acquitted ($x = 1$) it is given by $v(0) = -1/2$.

[13] The prosecution did not inform the defendant Brady of his companion's previous confession to the actual killing. The Supreme Court ruled that 'the government's withholding of evidence that is material to the determination of either guilt or punishment of a criminal defendant violates the defendant's constitutional right to due process.'

[14] This is in line with the broader development in continental Europe from medieval inquisitorial proceedings with secret charges to modern criminal law proceedings (Kittler, 2003). The most famous defendant whose charges are kept secret may be Josef K., the protagonist in Franz Kafka's novel *Der Proceß* (The Trial). In fact, Kittler (2003) suggests that legal scholar Kafka based his *Proceß* not on his contemporary but the medieval procedural standards.

guilty of the minor crime triggers a trial by misrepresenting his type as innocent he cannot hope for acquittal in the trial. The judge commits to punish her for having lied even if this is ex post inefficient $v(3/4) = 1/4 > 0$. This feature is reflected by the harsher punishments that lying in court or failing to confess usually entails.

2.4 The Proof of Theorem 2

2.4.1 The Case for Transparency

This subsection presents the main conceptual contribution of our paper: the principal cannot exploit the secrecy of a private signal that is correlated with the agent's type to reduce his information rents. This stands in marked contrast to settings with transfers where secrecy and correlation permit full elimination of information rents.

Proposition 1. *It is without loss of optimality for the principal to use a transparent procedure. Formally, consider any mechanism* $(\boldsymbol{x}, \boldsymbol{z})$ *that is feasible in the principal's problem. There exists a feasible mechanism* $(\tilde{\boldsymbol{x}}, \tilde{\boldsymbol{z}})$ *that satisfies* $(EPIC(s)_{t,\hat{t}})$ *for all* $s, t, \hat{t}$ *and delivers a payoff to the principal no lower than that generated by* $(\boldsymbol{x}, \boldsymbol{z})$.

The formal proof of this proposition can be found in the appendix. There are two channels through which the principal can potentially exploit secrecy. We revisit Example 1 to illustrate for each channel, (i) how it allows the principal to lower the verification costs required to implement an arbitrary allocation, and (ii) why this is not possible for optimal allocations.

Example 1 continued (a). *Consider the environment from Example 1 with distribution* (2.1)*. The optimal transparent verification schedule to implement the total allocation* (2.2)*, which allocates if and only if type and signal do not match, verifies with probability one whenever the good is allocated, i.e.*

$$\boldsymbol{z} = \left(\underset{(L,\ell)}{0}, \underset{(L,h)}{1}, \underset{(H,\ell)}{1}, \underset{(H,h)}{0} \right).$$

The agent's ex post incentive constraint is binding when his type and the signal match (and he does not get the good) and it is slack otherwise. Under secrecy, the total allocation (2.2) *can be implemented with the cheaper verification schedule*

$$\boldsymbol{z} = \left(\underset{(L,\ell)}{0}, \underset{(L,h)}{0.5}, \underset{(H,\ell)}{0.5}, \underset{(H,h)}{0} \right).$$

This creates only half the verification cost than with transparency.

With the second, cheaper verification schedule, the allocation is not transparently implementable. If type L knows that the signal is ℓ, he can get allocation with probability 1/2 by misreporting H. In this example, secrecy allows the principal to *re-use* excess allocation probability across different signals when the agent is unaware of the signal realisation. Why does this not work in the optimal allocation mechanism? When the total allocation probability is non-decreasing in t for all s, an improvement as above is not possible. First, monotonicity implies that only upward incentive constraints matter.[15] Second, if an EPIC constraint at a signal s binds for some type, it must also bind for all lower types at this signal. It follows that, in an optimal transparent mechanism, the lowest type's EPIC constraint must bind at all signals. No slack can be re-used even under secrecy. Note that the gain in flexibility to re-use excess utility across signals does not depend on correlation. This advantage of Bayesian Implementation is present for general distributions. In the case of independent types, however, it is precisely the pointwise monotonicity of allocations that leads to the EPIC-DIC equivalence (Manelli and Vincent, 2010, Gershkov et al., 2013, Ben-Porath et al., 2014).

With correlation, monotonicity in t is not sufficient to conclude that transparency is optimal. The resulting belief heterogeneity presents an additional channel to benefit from secrecy. The following example shows how the principal exploits this when the total allocation is non-monotone in s.

Example 1 continued (b). *Consider the total allocation probabilities*

$$\boldsymbol{x} + \boldsymbol{z} = \left(\underset{(L,\ell)}{1}, \underset{(L,h)}{0}, \underset{(H,\ell)}{1}, \underset{(H,h)}{1} \right). \tag{2.3}$$

The optimal verification probabilities to implement this transparently are

$$\boldsymbol{z} = \left(\underset{(L,\ell)}{0}, \underset{(L,h)}{0}, \underset{(H,\ell)}{0}, \underset{(H,h)}{1} \right).$$

Given the distribution in (2.1), *this results in verification costs of* $1 \cdot \frac{2}{6} c = \frac{1}{3} c$. *Without observing the signal, the agent updates his belief conditional on his type. Given the joint distribution in* (2.1), *type L's subjective belief on signal ℓ is* $\frac{2}{3}$. *Consider L's Bayesian incentive constraint:*

$$(BIC_{L,H}): \quad \frac{2}{3} \cdot 1 \quad + \quad \frac{1}{3} \cdot 0 \quad \geq \quad \frac{2}{3} \cdot (1 - z(H,\ell)) \quad + \quad \frac{1}{3} \cdot (1 - z(H,h)) \quad .$$

The principal can exploit that type L puts more weight on signal ℓ and shift verification probability from the type-signal combination (H,h) to (H,ℓ). Without trans-

[15]Indeed, the first step of the proof consists in presenting a relaxation of the principal's problem discarding, among others, all downward incentive constraints. Establishing monotonicity directly is complicated with belief heterogeneity.

parency, the allocation above is optimally implemented with verification probabilities

$$\boldsymbol{z} = \Big(\underset{(L,\ell)}{0}, \underset{(L,h)}{0}, \underset{(H,\ell)}{0.5}, \underset{(H,h)}{0}\Big). \tag{2.4}$$

They create verification costs of only $\frac{1}{6} \cdot 0.5 \cdot c = \frac{1}{12}c$.

This improvement is different from the one presented Example 1(a). It relies on correlation. The principal benefits from secrecy by exploiting the agent's belief heterogeneity, shifting verification probability from higher to lower signals. This reduces the overall verification probability because the relevant incentive constraint is for type L not to report H and type L's subjective belief on signal ℓ is higher. Why does this not work in the optimal mechanism? The reason that makes this shift possible when we move from transparency to secrecy is the non-monotonicity of the total allocation in the signal. In the proof we rule out this non-monotonicity. This is possible since the signal has no direct effect on the principal's preferences. Moving allocation probability towards higher signals for any given type relaxes the upward incentive constraints. As the signal s does not affect the principal's value directly, she is indifferent as to which signals carry allocation probability. This is different in our extension in which s has a direct effect on the principal's value (see Section 2.5).

However, pointwise monotonicity in both t and s is still not enough to obtain transparency. The following example shows that secrecy is also beneficial when allocation probabilities are interior.

Example 1 continued (c). *Consider the following total allocation probabilities:*

$$\boldsymbol{x} + \boldsymbol{z} = \Big(\underset{(L,\ell)}{0.5}, \underset{(L,h)}{0.5}, \underset{(H,\ell)}{1}, \underset{(H,h)}{1}\Big). \tag{2.5}$$

If the principal has to implement (2.5) *with a transparent procedure, the cost-minimal verification schedule is*

$$\boldsymbol{z} = \Big(\underset{(L,\ell)}{0}, \underset{(L,h)}{0}, \underset{(H,\ell)}{0.5}, \underset{(H,h)}{0.5}\Big).$$

This results in verification costs of $\frac{1}{6} \cdot 0.5\ c + \frac{2}{6} \cdot 0.5 \cdot c = \frac{1}{4} \cdot c$. *Just as in Example 1(b), the principal can exploit that type L puts more weight on signal* ℓ. *The optimal verification probabilities under secrecy,*

$$\boldsymbol{z} = \Big(\underset{(L,\ell)}{0}, \underset{(L,h)}{0}, \underset{(H,\ell)}{0.75}, \underset{(H,h)}{0}\Big),$$

ensure truthful reporting at lower costs $\frac{1}{6} \cdot 0.75 \cdot c = \frac{1}{8}\, c$.

Interior total allocation probabilities give the principal more flexibility in designing lie-deterring verification rules and allow her to benefit from secrecy. In the proof of Proposition 1, we show that it is optimal for the principal to use deterministic procedures. The optimality of transparent rules follows from the monotonicity and nonrandonmess of optimal allocation rules.

This concludes the intuition for the first step of the proof of Theorem 2. For clearer illustration, the chosen examples feature fixed total allocation probabilities and present changes in verification only. In the formal proof, however, we do not decompose the program, but solve for optimal allocation and verification jointly. This solution strategy is more effective in our setting as characterising optimal verification rules for arbitrary allocations is tedious due to the belief heterogeneity.

2.4.2 Optimal Transparent Mechanisms

To complete the proof of Theorem 2, we need to solve the principal's problem subject to the EPIC constraints.

Proposition 2. *The cutoff with appeal mechanism presented in Theorem 2 maximises the principal's payoff among all transparently implementable mechanisms.*

The formal proof is relegated to the appendix. Solving the problem under EPIC constraints is significantly simpler as the belief heterogeneity plays no role once the agent knows the signal. In fact, the principal's problem can be solved independently for each signal realisation s. The subproblem corresponding to signal s is analogous to the special case of a single agent in Ben-Porath et al. (2014). Hence, our main conceptual contribution, Proposition 1, creates a link between the cases of correlated and independent information by establishing that the principal voluntarily forgoes the screening potential of heterogeneous beliefs. This is in stark contrast to settings where monetary transfers are feasible.

2.5 Extension: When the Signal affects Preferences

In the previous sections, we characterise optimal mechanisms under the assumption that the principal's signal has no direct effect on her allocation value. In the court example in Section 2.3.1, the judge's value from convicting a guilty defendant does not depend on the findings from the pre-trial investigations. The relevance of these investigations stems from a better assessment of the defendant's guilt. *Given* his guilt or innocence, the investigation does not affect the judge's preferences over different rulings. Formally, the principal's valuation $v(t)$ does not depend on s directly.

To extend the analysis, suppose that the principal derives valuation $v : T \times S \to \mathbb{R}$ when allocating the good to an agent of type t at signal s. As before, we normalise the value she derives from not allocating to 0 and assume that $t \mapsto v(t,s)$ is nondecreasing for any $s \in S$. The agent's problem does not change. Both the Bayesian and EPIC incentive constraints remain the same. Only the principal's objective changes; it now reads:

$$\max_{(\boldsymbol{x},\boldsymbol{z})\geq 0} \mathbb{E}\left[\, x(t,s)\, v(t,s) + z(t,s)\, (v(t,s) - c)\,\right].$$

2.5.1 Optimal Transparent Mechanisms

We first characterise optimal transparent mechanisms. This characterisation and its proof are analogous to Proposition 2. The reason is that, by transparency, the optimal allocation and verification rule can again be determined separately for each signal realisation $s \in S$.

Proposition 3. *The optimal transparent mechanism is as follows: for all $s \in S$,*

$$\begin{cases} x(t,s) = 1,\ z(t,s) = 0 & \textit{if} \quad \mathbb{E}_t[\,(v(t,s) \mid s\,] > \mathbb{E}_t[\,(v(t,s) - c\,)^+ \mid s\,] \\ x(t,s) = 0,\ z(t,s) = \mathbb{1}_{\{v(t,s)>c\}} & \textit{otherwise.} \end{cases}$$

In addition, this characterisation leads to the following increasing (Figure 2.2(a)) and decreasing (Figure 2.2(b)) cutoff with appeals mechanisms under the respective regularity assumption:

(a) When the direct effect is positive, so that for all t, $s \mapsto v(t,s)$ is increasing, then there exists a cutoff $\bar{s}_a$ and appeal thresholds $\bar{t}(s)$ such that:

i) If $s \geq \bar{s}_a$, then $x(t,s) = 1$ for all t and $z(t,s) = 0$ for all t.

ii) If $s < \bar{s}_a$, then $x(t,s) = 0$ for all t and $z(t,s) = \begin{cases} 1 & \textit{for } t \geq \bar{t}(s) \\ 0 & \textit{for } t < \bar{t}(s), \end{cases}$

with $\bar{s}_a = \min\{s \mid \mathbb{E}_t[v(t,s)|s] > \mathbb{E}_t[(v(t,s)-c)^+ \mid s]\}$
and $\bar{t}(s) = \min\{t \mid v(t,s) - c > 0\}$.

(b) When the direct effect is sufficiently negative so that $s \mapsto \mathbb{E}_t[v(t,s)|s]$ is decreasing, then there exists a cutoff $\bar{s}_b$ and appeal thresholds $\bar{t}(s)$ such that:

i) If $s \geq \bar{s}_b$, then $x(t,s) = 0$ for all t and $z(t,s) = \begin{cases} 1 & \textit{for } t \geq \bar{t}(s) \\ 0 & \textit{for } t < \bar{t}(s). \end{cases}$

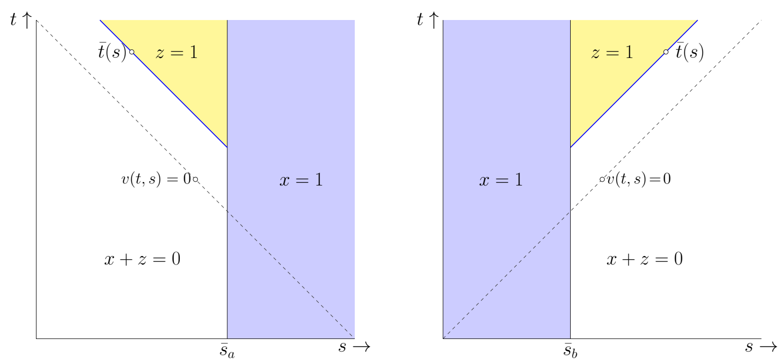
ii) If $s < \bar{s}_b$, then $x(t,s) = 1$ for all t and $z(t,s) = 0$ for all t,

with $\bar{s}_b = \min\{s \mid \mathbb{E}_t[v(t,s)|s] < \mathbb{E}_t[(v(t,s)-c)^+|s]\}$ *and* $\bar{t}(s)$ *as above.*

The formal proof in the appendix is analogous to the proof of Proposition 2. The general characterisation of the optimal transparent mechanism at the beginning of Proposition 3 holds true for any functional form of the principal's value $v(\cdot,\cdot)$. Under the monotonicity conditions in the second part of Proposition 3 the optimal transparent mechanism takes a cutoff with appeals form which is illustrated in Figure 2.2.

Figure 2.2(a) covers the case when the signal has a positive direct effect on the principal's value.[16] In the plot, $v(t,s) = t + s - 1/2$, so that the principal's first best allocation is to allocate whenever t exceeds the dotted decreasing diagonal. If the signal realisation is above the cutoff $\bar{s}_a$, the principal is optimistic about her value from allocating as v increases in s and a high s is indicative of a high type t. For $s \geq \bar{s}_a$ she allocates the good irrespective of his report without verification ($x = 1$). For signals below the cutoff $\bar{s}_a$, the good is allocated only after the agent's type is verified to be above the threshold $\bar{t}(s)$ which now depends on s. On both sides of the cutoff $\bar{s}_a$, this mechanism deviates from the principal's first best allocation to satisfy the agent's incentives. For $s \geq \bar{s}_a$, the good is allocated too often (the blue area below the dotted diagonal). For $s < \bar{s}_a$, the good is allocated too rarely (the white belt above the dotted diagonal).

Figure 2.2: Cutoff with appeals mechanisms with positive and negative direct effect



(a) Increasing $v(t,\cdot)$ (b) Sufficiently decreasing $v(t,\cdot)$

Figure 2.2(b) illustrates the cutoff mechanism when the direct effect is sufficiently negative so that $s \mapsto \mathbb{E}_t[v(t,s)|s]$ is decreasing.[17] The effect of s on $E_t[v(t,s)|s]$ consists of two components. There is a positive, indirect information effect: the principal expects a higher type after observing a higher signal. This is countered by a direct negative effect since $v(t,s)$ is decreasing in s. The condition in (b) means that the direct effect dominates. In the plot, $v(t,s) = t - s$, the principal's first best

[16] In the court example, this functional form would apply if the justice system got an additional benefit from the final verdict confirming the initial charge.

[17] Decreasing v in s would apply if the principal's signal increased her opportunity cost of allocating.

is to allocate whenever t exceeds the dotted increasing diagonal. Here, the principal is more optimistic about the allocation value when the signal is lower. Hence, the good is now allocated without verification when $s < \bar{s}_b$ and it is allocated only after verification when $s \geq \bar{s}_b$.

2.5.2 Optimal Bayesian Mechanisms

In this subsection, we show that our BIC-EPIC equivalence with correlation (Theorem 2) carries over when the direct effect is positive but not when it is negative. This dependence on the direction reveals the distinct nature of our setting in comparison to settings with independent information where the equivalence holds more generally.

If the direct effect is positive, the optimal Bayesian mechanism is again given by the optimal transparent mechanism.

Proposition 4. *If the direct effect of the signal is positive ($v(t, s)$ increasing in s), the transparent mechanism in Proposition 3(a) is optimal in the class of BIC mechanisms.*

The proof is relegated to the appendix. The optimal transparent mechanism in the case of a positive direct effect (Figure 2.2(a)) is deterministic and monotone in type and signal. Again, these properties ensure that no secret mechanism can do better than this transparent mechanism. The intuition that the transparent mechanism satisfies these properties is similar to the case without direct effect. In that case, in which the signal s did not affect the principal's value directly, she was indifferent as to which signals carry allocation probability. In the present case with a positive direct effect, $v(t, s)$ is increasing in s so that the principal prefers allocation probability at higher signals. Hence, the positive direct effect guarantees that the optimal transparent mechanism is increasing in s. This is no longer the case when the direct effect is negative. The example in Figure 2.2(b) shows that this is not the case if $v(t, s)$ is decreasing in s. Here, the optimal transparent mechanism is not increasing in s. In this case, the optimal BIC mechanism exploits secrecy and, thus, is not transparent:

Proposition 5. *If the direct effect of the signal is negative ($v(t, s)$ decreasing in s) and the optimal transparent mechanism $(\boldsymbol{x}, \boldsymbol{z})$ features $x(\hat{t}, s) > 0$ and $z(\hat{t}, s') > 0$ for some $\hat{t} \in T$ and some $s < s' \in S$, then there exists a mechanism $(\tilde{\boldsymbol{x}}, \tilde{\boldsymbol{z}})$ with a strictly higher value, which is BIC. Hence, the principal profits strictly from s being private.*

The proof is again relegated to the appendix. Note that under the regularity assumption in Proposition 3, that $\mathbb{E}_t[\ (v(T, s) \mid s\ ]$ is decreasing in s, the optimal transparent mechanism has the properties stated in the proposition whenever it is nontrivial, i.e. whenever $\boldsymbol{x}$ and $\boldsymbol{z}$ are positive at some combinations (t, s). As noted previously, the principal saves verification costs by shifting verification probability at high reports toward low signals. Low types, who have an incentive to misreport, find such signals more likely. Transparency then requires a shift in the allocation probability for low types toward the other signal realisations to ensure that, at those signals indicating no verification of high reports, the low types have no incentive to deviate. If the principal's value is decreasing in the signal, shifting allocation toward higher signals comes at a cost and, therefore, transparency comes at a cost.

2.6 Discussion and Concluding Remarks

2.6.1 Literature

In settings where monetary transfers are feasible, the principal can design lotteries rewarding the agent for guessing the value of her privately observed signal correctly. Different agent types hold different beliefs over the signal distribution and, therefore, reveal their type by guessing the signal they deem most likely. If the agent's liability is not limited, the principal can increase reward and loss in the lottery to such an extent that the incentives to win the lottery exceed any incentives regarding the allocation decision. In doing so, she can learn the agent's type at arbitrarily small costs. Mechanisms with monetary transfers and correlated information have been discussed by Crémer and McLean (1988), Riordan and Sappington (1988), Johnson et al. (1990), and McAfee and Reny (1992), who all establish conditions on the information structure that ensure full surplus extraction by the principal. As all surplus can be extracted, revenue maximisation leads to ex-post efficient allocations. Neeman (2004) discusses the genericity of the above-mentioned conditions and shows that full surplus extraction is possible only if every preference type is 'determined' by his belief over the correlated characteristics. Even though this condition is fulfilled in our setting with costly verification, full surplus extraction is not feasible and implementing the ex-post efficient allocation is not optimal for the principal. The full surplus-extracting lotteries require potentially unbounded transfers. For the case of bounded transfers or limited liability, Demougin and Garvie (1991) show that the qualitative results, the application of rewards as bets on the signal, still apply. Different from our setting, the principal gains by maintaining her signal private even when bounds or limited liability preclude full surplus extraction.

In the absence of monetary transfers, Bhargava et al. (2015) show how positively correlated beliefs among voters allow overcoming the impossibility of nondictatorial voting rules established by Gibbard (1973) and Satterthwaite (1975).

Our result is in line with the findings in other settings where monetary transfers are not feasible but correlated information is absent. The literature (Glazer and Rubinstein, 2004, Ben-Porath, Dekel and Lipman, 2019, Erlanson and Kleiner, 2020, Hart, Kremer and Perry, 2017, Halac and Yared, 2020) has found optimal mechanisms to take a simple cutoff structure and to be EPIC in the sense that the agents would also report truthfully if they were informed about the other agents' type realisations before their report.

The possibility for the mechanism designer to verify an agent's private information at a cost was first introduced by Townsend (1979) considering a principal-agent model for debt contracts, which was extended to a two-period model by Gale and Hellwig (1985). These early models of state verification feature both monetary transfers and verification. Glazer and Rubinstein (2004) introduce a setting where the principal has to take a binary decision depending on the multidimensional private information of the agent. Here, the principal cannot use monetary transfers, but she can learn about one dimension before making her decision.

Our model is most closely related to that of Ben-Porath, Dekel and Lipman (2014), who model costly verification and consider the case of allocating a good among finitely many agents whose types are independently distribute. Erlanson and Kleiner (2020) study a collective decision problem with costly verification and show that the optimal mechanism is EPIC and can be implemented by a simple weighted majority voting rule. Mylovanov and Zapechelnyuk (2017) consider an allocation problem without monetary transfers in which the principal learns the agents' types without cost but only posterior to the allocation decision and has the ability to punish untruthful reports up to a limit. Halac and Yared (2020) consider a delegation problem and specify conditions on the verification cost that ensure optimality of a threshold mechanism with an escape clause.

Erlanson and Kleiner (2020) show for the case of independent information that the equivalence between BIC and EPIC mechanisms holds more generally rather than only for optimal mechanisms. This relates to Gershkov, Goeree, Kushnir, Moldovanu and Shi (2013) and Manelli and Vincent (2010), who show equivalence between BIC and DIC mechanisms in settings with monetary transfers. All these results depend on the assumption that the private information of players is independently distributed (see discussion in Gershkov et al., 2013, p. 212). We deviate from this assumption by introducing correlation between the agent's type and the principal's signal and show that equivalence obtains for optimal allocation rules but not generally.

2.6.2 Informed principal problem

As the principal has private information, our model is also related to the informed principal problem cf. Myerson (1983) and Maskin and Tirole (1990). With monetary transfers, Severinov (2008) and Cella (2008) show that correlated information allows for an efficient solution to the informed principal problem. By our assumptions on the principal's commitment, the mechanism proposed by the designer does not convey information to the agent, so that there is no informed principal problem in our model Nonetheless, the fact that the principal's signal can be made public without loss implies that the informed principal game has a separating equilibrium in which the agent perfectly learns the principal's type from the proposal. This implies that our mechanism constitutes a solution to the informed principal problem for Cases 1 and 2 when the EPIC mechanism is optimal.

2.6.3 Concluding Remarks

This paper studies the role of information in a mechanism design model in which the principal may use costly verification instead of monetary transfers to incentivise the revelation of private information. We show that a transparent mechanism is optimal. It is without loss for the principal to make her information public before contracting with the agent. Our result gives a rationale for the use of transparent procedures in a variety of applications from hiring to procedural law. This is in contrast with results on correlation in mechanism design problems with money.

In an extension in which the principal's private information also affects her preferences, we characterise the mechanism and show that the above qualities remain if the information and direct effect work in the same direction. In the opposite case, we show how the principal can benefit by ensuring that her signal remains private.

2.A Appendix

Proof of Theorem 1

The revelation principle presented here is close to the revelation principle in Ben-Porath, Dekel and Lipman (2014), but it takes into account possible issues arising from the correlation between the signal and the type realisation.

Pick any (possibly dynamic) mechanism G and an agent strategy s_A that is a best response to this mechanism. Then, there is an equivalent incentive compatible, direct, two-stage mechanism characterised by the pair of functions (e, a),

$$e : T \times S \to [0, 1],$$
$$a : T \times T \cup \{\emptyset\} \times S \to [0, 1],$$

of the following form:

1. The agent reports his type $\hat{t} \in T$.
2. Given her signal realisation s, the principal verifies the agent's type with probability $e(\hat{t}, s)$.
3. Depending on the result of this verification $t \in T \cup \{\emptyset\}$, where $\emptyset$ encodes the event that there was no verification, the principal allocates the good to the agent with probability $a(\hat{t}, t, s)$.

Instead of G, the principal could commit to the following mechanism:

- The agent reports a type $\hat{t} \in T$.
- Given this report and her signal's realisation, s, the principal calculates the marginal probability of verification in the equilibrium in the original game under the condition that the agent's type was $\hat{t}$ and the principal's signal was s:[18]

$$e(\hat{t}, s) := \mathbb{P}(\text{there is verification}|s_A(\hat{t}), s).$$

- The principal verifies the agent's true type with this probability: $e(\hat{t}, s)$.
  - If she finds that the agent reported the truth, $\hat{t}$, or if she did not verify $t = \emptyset$, she allocates the good with probability that equals the marginal probability of allocation in the original mechanism, conditional on the type being equal to $\hat{t}$ and the signal being equal to s:
  $$a(\hat{t}, \hat{t}, s) = a(\hat{t}, \emptyset, s) = \mathbb{P}(\text{allocation}|s_A(\hat{t}), s).$$
  - If she verifies and finds out that the agent misreported, i.e. $t \notin \{\hat{t}, \emptyset\}$, the allocation probability is determined in the following way:
  The principal constructs a lottery over all stages in the original mechanism which have the principal verify the agent with positive probability in equilibrium, conditional on the event that the agent played according to $s_A(\hat{t})$ and the signal was s.
  The probabilities of the lottery are chosen such that they equal the probability of verifying at this stage for the first time, conditional on the event that there is verification at some point in the game.
  Now, she chooses one of these stages according to the above probabilities. She simulates the game from this point onward, assuming that the game had reached this stage and it was found at this point that the agent's true type was t, by letting the simulated agent behave according to what is described in $s_A(t)$ for behaviour after this knot and

[18] This means the probability that there was verification at any point in the game, specified by G and played by the agent according to $s_A(\hat{t})$, under the condition that signal s realised.

the verification. The strategy s_A contains a plan for the behaviour of the agent from this stage onward. The principal simulates his own behaviour, as prescribed in the original mechanism.

Given any signal realisation, this reproduces the allocation profile in the original game resulting from the following strategy for type $t \neq \hat{t}$ (which he could play without knowing the true signal realisation s): The agent of type t imitates type $\hat{t}$'s behaviour $s(\hat{t})$ until the first verification, and then sticks to the behaviour that the equilibrium strategy prescribes for his type.

If the agent reports the truth, the marginal probabilities of verification and allocation and, therefore, the expected utilities of the agent and the principal are the same in both mechanisms. However, truth-telling is optimal for the agent in the constructed mechanism, as misreporting yields the exact same outcome as the above-described deviation strategy in the original game and, therefore, cannot be profitable.

There are two further observations that help simplify the class of possible optimal mechanisms. In short, in any optimal mechanism, the principal will chose he highest possible punishment for detected misreports and the highest possible reward for detected truth-telling.

1. Maximal punishment: $t \notin \{\hat{t}, \emptyset\} \Rightarrow a(\hat{t}, t, s) = 0$
   As the mechanism is direct, in equilibrium, the agent will not lie, therefore, decreasing $a(\hat{t}, t, s)$ for $t \notin \{\hat{t}, \emptyset\}$ will not affect the expected utility of the mechanism designer. This deviation only increases the incentives to report truthfully. Therefore we can assume WLOG that the optimal mechanism features maximal punishment.
2. Maximal reward: $e(\hat{t}, s) > 0 \Rightarrow a(\hat{t}, \hat{t}, s) = 1$.
   Suppose $a(\hat{t}, \hat{t}, s) < 1$. One could now lower the probability of verification, $de(\hat{t}, s) < 0$, while increasing the probability of allocation after confirming the report as true, $da(s, \hat{t}, \hat{t}) > 0$, such that $d(e(\hat{t}, s)a(\hat{t}, \hat{t}, s)) = 0$.
   Lowering the verification probability would only increase the incentives to misreport and the overall allocation probability after report $\hat{t}$ and signal s, if there was allocation with positive probability conditional on no verification, i.e. $a(s, \hat{t}, \emptyset) > 0$. However, in this case, this allocation could be lowered $da(s, \hat{t}, \emptyset) < 0$ such that $d((1 - e(\hat{t}, s))a(s, \hat{t}, \emptyset)) = 0$, and the incentives to misreport and the overall allocation probability would remain constant. As this procedure would save verification costs while keeping all unconditional allocation probabilities constant, we can rule out that an optimal mechanism features non-maximal reward.

These observations fix the allocation after verification. Effectively the mechanism designer therefore has to choose only the verification probability $e(t, s)$ and the allocation probability, conditional on no verification $a(s, t, \emptyset)$.
For convenience, define $z(t, s) = e(t, s)$, the joint probability of verification and allocation, and $x(t, s) = (1 - e(t, s))a(t, \emptyset, s)$, the joint probability of no verification and allocation.
Note that the set of mechanisms described by

$$\{(x(t, s), z(t, s))_{t \in T, s \in S} | \forall t \in T \; \forall s \in S : \; 0 \leq x(t, s) + z(t, s) \leq 1\}$$

is equivalent to all maximal reward and punishment, two-stage, direct mechanisms.[19]

[19] The inverse mapping is given by

$$(e(t, s), a(t, \emptyset, s)) = (z(t, s), x(t, s)) \, .$$

Proof of Proposition 1

We present the following relaxation of the problem, and show that it is solved by a transparent mechanism which proves that it is a solution to the original LP. Define the set of profitable types as those t with a positive allocation value,

$$T^+ \equiv \{t \in T | v(t) > 0\},$$

and the unprofitable types accordingly as $T^- \equiv T \backslash T^+$. Both sets are non-empty by the assumption that v crosses 0. Otherwise, the optimal mechanism is trivial.

The relaxed problem includes only those incentive constraints that prevent types in T^- from misreporting types in T^+. Hence, it reads as follows:

$$(LP.r) \quad \max_{(\boldsymbol{x},\boldsymbol{z})\geq 0} \sum_{t\in T}\sum_{s\in S} f(t,s)\left[x(t,s)\ v(t) + z(t,s)\ (v(t)-c)\right]$$
$$\text{s.t.} \quad \forall t \in T^-,\ \forall \hat{t} \in T^+ : \quad (BIC_{t,\hat{t}}) \quad \text{and}$$
$$\forall (t,s) \in T \times S : \qquad x(t,s) + z(t,s) \leq 1.$$

In the remainder of the proof, we derive feasible changes to a solution to the relaxed problem which do not lower the principal's value and which finally lead to the cutoff mechanism. We make repeated use of the following notation: we denote changes in the allocation probability by $dx(t,s)$ so that the new probability after the change is given by $x(t,s) + dx(t,s)$. $dx(t,s)$ may be positive or negative. Analogously for $dz(t,s)$. Further, $d(BIC_{t,\hat{t}})$ denotes the change in surplus utility that type t receives from reporting the truth rather than misreporting $\hat{t}$, which is induced by a change of the above form. Recall that the constraint $(BIC_{t,\hat{t}})$ reads as $\sum_s f(t,s)\left[x(t,s) + z(t,s) - x(\hat{t},s)\right] \geq 0$ so that $d(BIC_{t,\hat{t}})$ denotes the change to the left-hand side of this inequality.

The value for the principal is given by

$$V = \sum_{t\in T}\sum_{s\in S} f(t,s)\left[x(t,s)\ v(t) + z(t,s)\ (v(t)-c)\right],$$

and dV will denote the induced change to this value.

Step 1: The optimal mechanism in the relaxed problem features $\forall t \in T^-\ \forall s \in S:\ z(t,s) = 0$:

Suppose $z(t,s) > 0$ for some type $t \in T^-$. Shifting probability mass from $z(t,s)$ to $x(t,s)$ such that the overall allocation probability stays constant,

$$0 < dx(t,s) = -dz(t,s),$$

saves the principal verification costs and does not distort the incentives, as type t's incentive to misreport remains the same, and all incentive constraints to misreport a type $t \in T^-$ are ignored in the relaxed problem.

Step 2: There is an optimal mechanism in the relaxed problem featuring a cutoff form for $\boldsymbol{x}(\hat{t},\cdot)$:

$$\forall \hat{t} \in T^+\ \exists \tilde{s}(\hat{t}) \in S:\ x(\hat{t},s) \begin{cases} = 0 & \text{if } s < \tilde{s}(\hat{t}) \\ \in [0,1) & \text{if } s = \tilde{s}(\hat{t}) \\ = 1 & \text{if } s > \tilde{s}(\hat{t}) \end{cases}.$$

Take a feasible IC mechanism of the relaxed problem featuring that for some $\hat{t} \in T^+$, $\exists s < s' \in S$ such that $x(\hat{t},s) > 0$, $x(\hat{t},s') < 1$.

Note that the value of $a(t,\emptyset,s)$ does not play any role in the mechanism if $e(t,s) = z(t,s) = 1$ and can therefore be chosen arbitrarily.

Modify the mechanism only at two points, shifting allocation probability mass from $x(\hat{t}, s)$ to $x(\hat{t}, s')$, i.e. $dx(\hat{t}, s) < 0$ and $dx(\hat{t}, s') > 0$. Choose these shifts in a proportion such that for the highest unprofitable type, $\tilde{t} \equiv \max T^-$, the incentive to misreport $\hat{t}$ remains unchanged:

$$0 \overset{!}{=} d(BIC_{\tilde{t},\hat{t}}) = -f(\tilde{t}, s)\, dx(\hat{t}, s) - f(\tilde{t}, s')\, dx(\hat{t}, s') = 0 \Leftrightarrow dx(\hat{t}, s) = -\frac{f(\tilde{t}, s')}{f(\tilde{t}, s)}\, dx(\hat{t}, s').$$

For all types $t \in T^-$, we have $t \leq \tilde{t}$, and, therefore,

$$d(BIC_{t,\hat{t}}) = -f(t, s)\, dx(\hat{t}, s) - f(t, s')\, dx(\hat{t}, s') = f(t, s) \left[\frac{f(\tilde{t}, s')}{f(\tilde{t}, s)} - \frac{f(t, s')}{f(t, s)}\right] dx(\hat{t}, s') \geq 0$$

by the monotone likelihood ratio property. The principal's value changes in the following way:

$$\begin{aligned} dV =& f(\hat{t}, s)\, dx(\hat{t}, s)\, v(\hat{t}) + f(\hat{t}, s')\, dx(\hat{t}, s')\, v(\hat{t}) \\ =& f(\hat{t}, s) \left[-\frac{f(\tilde{t}, s')}{f(\tilde{t}, s)}\, dx(\hat{t}, s')\right] v(\hat{t}) + f(\hat{t}, s')\, dx(\hat{t}, s')\, v(\hat{t}) \\ =& f(\hat{t}, s) \left[\frac{f(\hat{t}, s')}{f(\hat{t}, s)} - \frac{f(\tilde{t}, s')}{f(\tilde{t}, s)}\right] dx(\hat{t}, s')\, v(\hat{t}) \geq 0, \end{aligned}$$

since $dx(\hat{t}, s') > 0$ and $\hat{t} \in T^+$, which implies both $v(\hat{t}) \geq 0$ and $\hat{t} > \tilde{t}$. The proposed shift is clearly feasible if in the original mechanism, $x(\hat{t}, s') + z(\hat{t}, s)' < 1$. In the case that $x(\hat{t}, s') + z(\hat{t}, s)' = 1$, it can still be implemented by shifting in addition mass from $z(\hat{t}, s')$ to $z(\hat{t}, s)$ to ensure that $x(\hat{t}, s') + z(\hat{t}, s)'$ and $x(\hat{t}, s) + z(\hat{t}, s)$ remain constant:

$$dx(\hat{t}, s') + dz(\hat{t}, s') = 0 \;\text{ and }\; dx(\hat{t}, s) + dz(\hat{t}, s) = 0.$$

This implies $dz(\hat{t}, s') < 0$ and $dz(\hat{t}, s) > 0$. This is feasible as $x(\hat{t}, s') < 1$ and $x(\hat{t}, s') + z(\hat{t}, s)' = 1$ imply that $z(\hat{t}, s') > 0$. As $x(\hat{t}, s) > 0$, we must further have $z(\hat{t}, s) < 1$ by feasibility. To maintain the total allocation probabilities constant, the above changes in x are compensated by the following changes in z:

$$dz(\hat{t}, s) = \frac{f(\tilde{t}, s')}{f(\tilde{t}, s)}(-dz(\hat{t}, s')).$$

The incentives for any lower type to misreport his type as $\hat{t}$ are weakened in the same way as above because $z(\hat{t}, s)$ and $z(\hat{t}, s')$ do not play a role in the constraints that prevent misreport $\hat{t}$.

Finally, the principal's value now changes by

$$\begin{aligned} dV =& f(\hat{t}, s) \left[dx(\hat{t}, s)\, v(\hat{t}) + dz(\hat{t}, s)\, (v(\hat{t}) - c)\right] + f(\hat{t}, s') \left[dx(\hat{t}, s')\, v(\hat{t}) + dz(\hat{t}, s')\, (v(\hat{t}) - c)\right] \\ =& -c \left[f(\hat{t}, s)\, dz(\hat{t}, s) + f(\hat{t}, s')\, dz(\hat{t}, s')\right] \\ =& -c\, f(\hat{t}, s) \left[\frac{f(\tilde{t}, s')}{f(\tilde{t}, s)} - \frac{f(\hat{t}, s')}{f(\hat{t}, s)}\right] (-dz(\hat{t}, s')) \geq 0, \end{aligned}$$

as, by MLRP, the term in squared brackets is negative and, by assumption, $-dz(\hat{t}, s') \geq 0$.

Step 3: There is an optimal mechanism in the relaxed problem featuring $\boldsymbol{x}(\hat{t}, \cdot) = \boldsymbol{x}(\hat{\hat{t}}, \cdot)$ for all $\hat{t}, \hat{\hat{t}} \in T^+$:

By the cutoff structure established in Step 2, $\boldsymbol{x}(\hat{t}, \cdot) = (0, \ldots, 0, x(\hat{t}, \tilde{s}(\hat{t})), 1, \ldots, 1)$ for all $t \in T^+$. Suppose to the contrary that $x(\hat{t}, \tilde{s}(\hat{t})) + \sum_{s > \tilde{s}(\hat{t})} 1 > x(\hat{\hat{t}}, \tilde{s}(\hat{\hat{t}})) + \sum_{s > \tilde{s}(\hat{\hat{t}})} 1$ for some $\hat{t}, \hat{\hat{t}} \in T^+$. Replacing $\boldsymbol{x}(\hat{\hat{t}}, \cdot)$ by $\boldsymbol{x}(\hat{t}, \cdot)$ does not generate new incentives to misreport, but it increases the principal's expected value, as it increases the allocation probability for profitable types. If feasibility is hurt, i.e. $x(\hat{t}, s) + z(\hat{\hat{t}}, s) > 1$ for some $s \in S$, decrease $z(\hat{t}, s)$ until $x(\hat{t}, s) + z(\hat{\hat{t}}, s) = 1$. This is also a strict improvement for the principal, as she saves verification costs.

Step 4: There is an optimal mechanism in the relaxed problem featuring $\boldsymbol{x}(\hat{t},\cdot) = \boldsymbol{x}(\hat{\hat{t}},\cdot)$ for all $\hat{t},\hat{\hat{t}} \in T^+ \cup T^-$:

Fix some unprofitable type $t \in T^-$. By Step 1, we have $\boldsymbol{z}(t,\cdot) = 0$. Optimally, the principal wants to choose the lowest possible allocation probability for the unprofitable types. However, she needs to grant him at least the same interim allocation probability that he could achieve by misreporting to be a profitable type $\hat{t} \in T^+$ (by steps 2–3, we know that $\boldsymbol{x}(\hat{t},\cdot)$ is the same for all $\hat{t} \in T^+$). As the signal realisation has no effect on the allocation value, the principal is indifferent between any allocation vector $\boldsymbol{x}(t,\cdot)$ which induces the same interim allocation probability, $\mathbb{E}[x(t,s)|t]$, i.e. $\mathbb{E}[v(t)x(t,s)|t] = v(t)\mathbb{E}[x(t,s)|t]$. Therefore, she can grant the unprofitable types just the same allocation probability they would face if they misreported a profitable type: $\boldsymbol{x}(t,\cdot) = \boldsymbol{x}(\hat{t},\cdot)$.

This step concludes the proof. If the non-verified allocation probability is independent of the type report at each signal, the agent has no incentive to misreport even if he knows the signal.

□

Proof of Proposition 2

Step 0: For any $s \in S$, the optimal $(\boldsymbol{x}(\cdot,s),\boldsymbol{z}(\cdot,s))$ can be determined separately, as all constraints only involve allocation and verification probabilities for the same signal realisation. The principal's optimal expected value is the weighted sum of the values of these subproblems:

$$(LP(s)) \qquad \max_{(\boldsymbol{x}(\cdot,s),\boldsymbol{z}(\cdot,s))\geq 0} \mathbb{E}_t \left[\, x(t,s)\; v(t) \;+ z(t,s)\; (v(t)-c) \mid s \,\right]$$

$$\text{s.t. } \forall t,\hat{t} \in T: \; (EPIC(s)_{t,\hat{t}}) \quad \text{and}$$

$$\forall t \in T: \qquad x(t,s) + z(t,s) \leq 1.$$

Step 1: *For any $s \in S$ and for all $t,\hat{t} \in T$: $x(t,s) = x(\hat{t},s)$, i.e. the allocation probability $\boldsymbol{x}(\cdot,s)$ has to be constant in the report.*

Suppose to the contrary that there were reports t and $\hat{t}$ with $x(\hat{t},s) > x(t,s)$. Ex-post incentive compatibility implies that for all $\tilde{t} \in T$, we have $x(\tilde{t},s) + z(\tilde{t},s) \geq x(\hat{t},s) > x(t,s)$. Hence, there cannot be a type with a binding incentive constraint regarding the report t. This, in turn, implies that optimally, $z(t,s) = 0$. If it were positive, $z(t,s)$ could be lowered and $x(t,s)$ could be increased, at least until the strict inequality above binds. This leaves the allocation probabilities unchanged but lowers verification costs.

The incentive constraints of type t now take the form $x(t,s) + 0 \geq x(\tilde{t},s)$ for all reports $\tilde{t}$ and, in particular, for report $\hat{t}$, contradicting the above hypothesis. Hence, we must have that for all t, $\hat{t}$: $x(t,s) = x(\hat{t},s) \equiv \chi_s$.

Step 2: With constant $\boldsymbol{x}(\cdot,s)$, all incentive constraints are automatically fulfilled, as the unverified allocation probability is the same for any possible report. The principal's problem reads as follows:

$$(LP(s)) \qquad \max_{(\chi_s,\boldsymbol{z}(\cdot,s))\geq 0} \sum_{t\in T} f(t,s)\left[\chi_s\; v(t) + z(t,s)\; (v(t)-c)\right]$$

$$\text{s.t. } \quad \forall t \in T: \chi_s + z(t,s) \leq 1.$$

In this simplified program, $z(t,s)$ will be set as high as possible, i.e. to $1-\chi_s$ if $(v(t,s)-c)$ is positive and to 0 otherwise, yielding the following:

$$(LP(s)) \quad \max_{\chi_s\in[0,1]} \quad \chi_s \cdot \sum_{t\in T} f(t,s)\; v(t) \quad + \quad \sum_{t\in T} f(t,s)\; (1-\chi_s)\; (v(t)-c)^+.$$

Step 3: Expressed in terms of conditional expectations, the problem is linear in χ_s:

$$(LP(s)) \quad \max_{\chi_s \in [0,1]} \; \chi_s \cdot \mathbb{E}_t\big[v(t)|\, s\,\big] + (1-\chi_s)\cdot \mathbb{E}_t\big[(v(t)-c)^+|\, s\,\big].$$

Generically, the optimal value of χ is either 0 or 1, depending on which of the expectations is larger. The optimality of the cutoff mechanism follows as $\mathbb{E}_t[\, v(t) \mid s\,] - \mathbb{E}_t[\,(v(t)-c\,)^+ \mid s\,]$ is increasing in s. □

Proof of Proposition 3

Similar to Step 0 in Proposition 2, for any $s \in S$, the optimal $(\boldsymbol{x}(\cdot,s), \boldsymbol{z}(\cdot,s))$ can be determined separately, as all constraints only involve allocation and verification probabilities for the same signal realisation. This results in $|S|$ separate problems, one for each possible signal realisation $s \in S$. In these separate problems $v(t,s)$ is a function of t only. Therefore, for all $s \in S$, all steps in the proof of Proposition 2 can be replicated with $v(t)$ replaced by $v(t,s)$.

The results on the cutoff mechanisms follows from the above solution for each s and the observation that the regularity assumptions imply that

$$\mathbb{E}_t[\, v(t,s) \mid s\,] - \mathbb{E}_t[\,(v(t,s)-c\,)^+ \mid s\,]$$

is monotone in s; increasing in case (a) and decreasing in case (b). □

Proof of Proposition 4

Step 0: (Relaxation) Define the following cutoff in the signal space:

$$\overline{s} = \min\{s|\mathbb{E}[v(t,s)|s] > \mathbb{E}[(v(t,s)-c)^+]\},$$

where $(a)^+ = \max\{0,a\}$ and we use the convention that $\min\emptyset = \max S$. Note that $v(t,s) - (v(t,s)-c)^+ = \min\{c, v(t,s)\}$ is increasing in both components. Due to the MLRP it follows that $E[v(t,s)|s] - \mathbb{E}[(v(t,s)-c)^+$ is increasing in s.

Next, define the set of profitable types $T^+ = \{t \in T|\, v(t,\overline{s}) \geq 0\}$. We denote all types that are not profitable by $T^- = T - T^+$.

The relaxed problem ignores certain incentive constraints. It optimises the same objective function but only subject to:

$$\forall \hat{t} \in T^+ \; \forall t \in T \text{ with } t < \hat{t}\text{: } (BIC_{t,\hat{t}})$$

Step 1: There is an optimal solution to the relaxed problem that takes a cutoff form in x for all $t \in T$:

$$\forall t \in T \; \exists\, \tilde{s}(t) \in S: \; x(t,s) = \begin{cases} 0 & \text{if } s < \tilde{s}(t) \\ x(t,s) \in [0,1) & \text{if } s = \tilde{s}(t) \\ 1 & \text{if } s > \tilde{s}(t) \end{cases}.$$

Suppose there is a (relaxed) incentive compatible mechanism which has for some $t \in T$ and $s' < s'' \in S$: $x(t,s) > 0$ and $x(t,s') < 1$.

In the following, we consider a shift in allocation probability from $x(t,s')$ to $x(t,s'')$ that keeps the overall allocation probability for type t constant:

$$f(t,s')\, dx(t,s') + f(t,s'')\, dx(t,s'') = 0 \Leftrightarrow \underbrace{dx(t,s')}_{<0} = -\frac{f(t,s'')}{f(t,s')}\, \underbrace{dx(t,s'')}_{>0}.$$

First, for any type $t^- < t$ the probability of receiving the good without verification after a misreport t decreases in the new mechanism:

$$f(\tilde{t},s')dx(t,s') + f(\tilde{t},s'')dx(t,s'') = -f(\tilde{t},s')\left[\frac{f(t,s'')}{f(t,s')} - \frac{f(\tilde{t},s'')}{f(\tilde{t},s')}\right]dx(t,s'') \leq 0.$$

The last inequality holds since the likelihood ratio is increasing. The shift yields type t the same allocation probability, so he cannot have a new incentive to misreport. Therefore, all relaxed incentive constraints survive.

Second, the modified mechanism yields the principal a higher expected value:

$$f(t,s')\;dx(t,s')\;v(t,s') + f(t,s'')\;dx(t,s'')\;v(t,s'') = f(t,s'')dx(t,s'')\left[-v(t,s') + v(t,s'')\right] > 0$$

The proposed shift is clearly feasible if in the original mechanism $x(t,s'') + z(t,s'') < 1$. In the case that $x(t,s'') + z(t,s'') = 1$, it can still be implemented by shifting in addition mass from $z(t,s'')$ to $z(t,s')$ such that $x(t,s'') + z(t,s'')$ and $x(t,s') + z(t,s')$ stay constant:

$$dx(t,s') = -dz(t,s') \;\text{ and }\; dx(t,s'') = -dz(t,s'').$$

As we assume $x(t,s'') < 1$ and $x(t,s'') + z(t,s'') = 1$, we have $z(t,s') > 0$. Since $x(t,s') > 0$ we also have $z(t,s') < 1$.

The incentives for any lower type to misreport his type as t are weakened in the same way as above since $z(t,s)$ and $z(t,s')$ do not play a role in these constraints. The incentive for t to misreport is not affected since the total allocation probability $x + z$ is kept constant.

Further, the principal's expected value is not changed by these shifts:

$$\begin{aligned} f(t,s')\left[dx(t,s')\;v(t,s') + dz(t,s')\;(v(t,s') - c)\right] + f(t,s'')\left[dx(t,s'')\;v(t,s'') + dz(t,s'')\;(v(t,s'') - c)\right] \\ = -c\;f(t,s'')\left[\frac{f(t,s'')}{f(t,s')} - \frac{f(t,s'')}{f(t,s')}\right](-dz(t,s')) = 0. \end{aligned}$$

The reason is that the allocation probability $x + z$ remains the same with these shifts, so that only the verification changes. However, the change in verification is such that it does not alter the expected verification probability for the true type t and, therefore, neither the expected verification cost for the principal.

Step 2: The optimal mechanism in the relaxed problem features

$$\forall t \in T^- \;\forall s \in S: \; z(t,s) = 0.$$

In the relaxed problem we disregard all incentive constraints that prevent the agent to misreport his type as $t \in T^-$. If there were some $t \in T^-$ and $s \in S$ with $z(t,s) > 0$, shifting probability mass from $z(t,s)$ to $x(t,s)$ by

$$\underbrace{dz(t,s)}_{<0} = -\underbrace{dx(t,s)}_{>0},$$

would save the principal verification costs while keeping the overall allocation probability constant. It would therefore not affect the incentive constraints in the relaxed problem.

Step 3: We can assume that the optimal mechanism also takes a cutoff form in $x + z$:

$$\forall t \in T\;\exists \underline{s}(t) \in S: \; x(t,s) + z(t,s) = \begin{cases} 0 & \text{if } s < \underline{s}(t) \\ x(t,s) + z(t,s) \in [0,1) & \text{if } s = \underline{s}(t) \\ 1 & \text{if } s > \underline{s}(t) \end{cases}.$$

For $t \in T^-$, this property follows immediately from the previous two steps with $\underline{s}(t) = \tilde{s}(t)$.

Suppose for $t \in T^+$ that there exist $s' < s'' \in S$ with $z(t,s') > 0$ and $x(t,s'') + z(t,s'') < 1$. To rule out this possibility, consider a shift in mass from $z(t,s')$ to $z(t,s'')$ in a way that the allocation probability for a truth-telling agent of type t remains constant, i.e.

$$\underbrace{dz(t,s'')}_{>0} = \frac{f(t,s')}{f(t,s'')}(-\underbrace{dz(t,s')}_{<0}).$$

Note that this shift is feasible by assumption and that it will keep all relaxed incentive constraints unchanged, since the true type t receives the same expected allocation probability, and $\boldsymbol{z}(t,\cdot)$ does not play a role in the IC constraints preventing misreport t.

From the principal's point of view, it is favourable because it keeps the verification probability and thus the costs constant, while shifting allocation mass from (t,s') to the more favourable type–signal pair (t,s''), i.e.

$$\begin{aligned} dV &= f(t,s')dz(t,s')[v(t,s') - c] + f(t,s'')dz(t,s'')[v(t,s'') - c] \\ &= 0 \cdot c + f(t,s')[v(t,s'') - v(t,s')](-dz(t,s')) > 0. \end{aligned}$$

Step 4: In the relaxed problem, it is without loss for the principal to require the IC constraints to hold point-wise at each signal, i.e. $\forall t \in T$, $\forall \hat{t} \in T^+$ with $t < \hat{t}$ and $\forall s \in S$:

$$(EPIC(s)_{t,\hat{t}}) : x(t,s) + z(t,s) - x(\hat{t},s) \geq 0.$$

By the above steps, the (Bayesian) IC constraints in the relaxed problem can be written as follows:[20] $\forall t \in T$ and $\forall \hat{t} \in T^+$ with $\hat{t} > t$:

$$\begin{aligned} &\sum_{s \in S} f(t,s)(x(t,s) + z(t,s)) && - \sum_{s \in S} f(t,s)x(\hat{t},s) \\ =& f(t,\underline{s}(t))(x(t,\underline{s}(t)) + z(t,\underline{s}(t))) + \sum_{s > \underline{s}(t)} f(t,s)\, 1 && -(f(t,\tilde{s}(\hat{t}))\, x(\hat{t},\tilde{s}(\hat{t})) + \sum_{s > \tilde{s}(\hat{t})} f(t,s)\, 1) \geq 0. \end{aligned}$$

This condition clearly requires that $\underline{s}(t) \leq \tilde{s}(\hat{t})$ and, in the case of equality, $x(t,\underline{s}(t)) + z(t,\underline{s}(t)) \geq x(\hat{t},\tilde{s}(t))$. Because by the definition of $\underline{s}(t)$, $x + z$ is equal to 0 below and equal to 1 above this threshold, we can conclude that, for all s, $x(t,s) + z(t,s) \geq x(\hat{t},s)$ which implies $(EPIC(s)_{t,\hat{t}})$.

Step 5: Consider an optimal mechanism in the relaxed problem that satisfies the cutoff structure from the previous step. This mechanism also satisfies $(EPIC(s)_{t,\hat{t}})$ for all $t,\hat{t} \in T^-$ with $t < \hat{t}$. That is, no unprofitable type has an incentive to report any higher unprofitable type.

Assume that for some $s \in S$ there are types $t < \hat{t} \in T^-$ such that the constraint $(EPIC(s)_{t,\hat{t}})$ is violated. Define $s' \equiv \min\{s \in S | \exists t < \hat{t} \in T^- : x(t,s) + z(t,s) < x(\hat{t},s)\}$ to be the lowest signal for which some type t profits from a higher report $\hat{t} \in T^-$. Let $t' \equiv \min\{t \in T^- | \exists \hat{t} \in T \text{ with } \hat{t} > t : x(t',s') < x(\hat{t},s')\}$ be the smallest type with $EPIC(s')$ incentives to misreport his type to some type $\hat{t} \in T^-$.

Since $z(t',s') = 0$ for the unprofitable type t' (Step 3), this implies $x(t,s') < x(\hat{t},s')$. As $t',\hat{t} \in T^-$, it follows that $\hat{t}$'s EPIC(s') constraints are slack for all reports in T^+. Having $x(\hat{t},s') > x(t',s') \geq 0$ can therefore only be optimal in the relaxed problem if $v(\hat{t},s') \geq 0$. This implies that $s' > \bar{s}$ since T^- is precisely defined as the set of types t with $v(t,\bar{s}) < 0$. Since $x(t',s') < x(\hat{t},s') \leq 1$, taking the cutoff structure from step 1 into account we can infer that for all $s < s'$ it holds that $x(t',s') = 0$. In particular we have $x(t',\bar{s}) = 0$.

By the minimality of s' we get that $0 = x(t',\bar{s}) \geq x(\tilde{t},\bar{s})$ for all $\tilde{t} \in T^-$ with $\tilde{t} > t'$. By the

[20] Making use of the fact that $z(t,s) = 0$ for all $t \in T^-$.

minimality of t' we get that $x(\tilde{t}, \bar{s}) \leq x(t', \bar{s}) = 0$ for all $\tilde{t} \in T^-$ with $\tilde{t} < t'$. Furthermore, by $EPIC(\bar{s})$ IC compatibility, it follows that $0 = x(t, \bar{s}) \geq x(t', \bar{s})$ for all $t' \in T^+$. So we have that $x(t, \bar{s}) = 0$ for all t which by definition of $\bar{s}$ cannot be optimal in the restricted problem (for which this mechanism is also optimal) as $\mathbb{E}_t[\, (v(t, \bar{s}) \mid \bar{s}\,] > \mathbb{E}_t[\, (v(t, \bar{s}) - c\,)^+ \mid \bar{s}\,] \geq 0$.

Step 6: Consider an optimal mechanism in the relaxed problem that satisfies $(EPIC(s)_{t,\hat{t}})$ for all $t \in T$, for all $\hat{t} \in T^+$ with $\hat{t} > t$, and for all $s \in S$. This mechanism also satisfies $(EPIC(s)_{t,\hat{t}})$, for $t, \hat{t} \in T$ with $t > \hat{t}$, i.e. no type benefits from reporting any lower type.

Assume that there are types $t'' > t' \in T$ such that $x(t'', s) + z(t'', s) < x(t', s)$ for some s. WLOG let t'' be the lowest type for which such a downward deviation is profitable.

Optimality of the relaxed mechanism requires then that $\mathbb{E}_t[\, v(t, s)\mathbb{1}_{\{T \leq t'\}} \mid s\,] > 0$. Otherwise the principal would be better off by lowering x for all types below t'' (note that $x(t'', s) + z(t'', s) < x(t', s)$ implies that types $t < t''$ cannot have binding upwards constraints towards reports higher than t'' as this would violate the upward constraints for t''). Monotonicity of the value in the type in turn implies that $v(t, s) > 0$ for all $t > t'$. This contradicts optimality as the designer could increase $x(t, s)$ for all higher types without violating any incentives. Either by just increasing $x(t, s)$ if $x(t, s) + z(t, s) < 1$ or by lowering $z(t, s)$ at the same time to save verification costs.

This concludes the proof. We have shown that optimal solution to the relaxed problem is EPIC (Step 4). Therefore the principal's expected value in the original problem cannot exceed the expected value from this optimal EPIC solution. In Steps 5 and 6 we ruled out the two possible violations of the original (ex-post) incentive constraints that can arise in a solution to the relaxed problem. Hence the candidate solution is also EPIC in the original problem. In particular, it is Bayesian incentive compatible and therefore a solution to the original problem. □

Proof of Proposition 5

To prove the claim, we construct an improvement that will not violate the Bayesian incentive constraints. This suffices to show that the principal strictly profits from ensuring that the realisation of s remains private because the improved mechanism will implement the same allocation at lower verification costs. Consider the shift of mass from $z(\hat{t}, s')$ to $z(\hat{t}, s)$ and – in order to maintain the overall allocation $\boldsymbol{x} + \boldsymbol{z}$ unchanged – vice versa for $x(\hat{t}, s')$ and $x(\hat{t}, s)$:

$$dx(\hat{t}, s') + dz(\hat{t}, s') = 0 \text{ and } dx(\hat{t}, s) + dz(\hat{t}, s) = 0.$$

To ensure that the Bayesian incentive constraints of all types $t < \hat{t}$ are not violated by the shift, we require that

$$\forall t < \hat{t}: \; d(BIC_{t,\hat{t}}) = -f(t, s)\, dx(\hat{t}, s) - f(t, s')\, dx(\hat{t}, s') \geq 0,$$

which is equivalent to $dx(\hat{t}, s') \leq \frac{f(t,s)}{f(t,s')} \left(-dx(\hat{t}, s)\right)$. The proposed change has $-dx(\hat{t}, s) > 0$, and $\frac{f(t,s)}{f(t,s')}$ is decreasing in t. Hence, the right-hand side of the above expression is minimised at $t' = \max\{t \in T | t < \hat{t}\}$. Note that $\hat{t} \neq \min\{t \in T\}$ as otherwise $v(t, s') > c > 0$ at all t so that the optimal mechanism would allocate without verification after this signal.

Setting $dx(\hat{t}, s') = \frac{f(t',s)}{f(t',s')} \left(-dx(\hat{t}, s)\right)$ ensures that the incentives to misreport toward $\hat{t}$ are weakened for all lower types. The above changes in x imply for z:

$$-dz(\hat{t}, s) = \frac{f(t', s')}{f(t', s)} dz(\hat{t}, s').$$

The principal's value changes as follows:

$$\begin{aligned} dV =& f(\hat{t},s)\left[dx(\hat{t},s)\ v(\hat{t},s) + dz(\hat{t},s)\ (v(\hat{t},s)-c)\right] + f(\hat{t},s')\left[dx(\hat{t},s')\ v(\hat{t},s') + dz(\hat{t},s')\ (v(\hat{t},s')-c)\right] \\ =& -c\left[f(\hat{t},s)\ dz(\hat{t},s) + f(\hat{t},s')\ dz(\hat{t},s')\right] \\ =& -c\ f(\hat{t},s)\left[\frac{f(t',s')}{f(t',s)} - \frac{f(\hat{t},s')}{f(\hat{t},s)}\right](-dz(\hat{t},s')) > 0. \end{aligned}$$

The second equality follows because the allocation remains the same with these shifts, so that only the verification cost changes.

Finally, note that in the optimal EPIC mechanism, $z(\hat{t},s)=1$ implies $z(t,s)=1$ for all $t>\hat{t}$ and that $\boldsymbol{x}(\cdot,s)$ is constant in the report at all s. Therefore, the fact that $z(\hat{t},s)=1$ in the original mechanism implies that the Bayesian IC constraints for higher types to lie downward to $\hat{t}$ are slack so that we can always find a shift in magnitude small enough to not violate these constraints. The only case in which these constraints are not slack in the optimal EPIC mechanism is when several types receive exactly the same allocation. In this case, the above improvement can be applied to the highest report in this class. □

References

Akbarpour, M. and Li, S. (2020). Credible Auctions: A Trilemma, *Econometrica* **88**(2): 425–467.

Austria (n.d.). Die österreichische Strafprozessordnung - StPO.
URL: *www.jusline.at/gesetz/stpo/paragraf/6*

Ben-Porath, E., Dekel, E. and Lipman, B. L. (2014). Optimal Allocation with Costly Verification, *American Economic Review* **104**(12): 3779–3813.

Ben-Porath, E., Dekel, E. and Lipman, B. L. (2019). Mechanisms with Evidence: Commitment and Robustness, *Econometrica* **87**(2): 529–566.

Bhargava, M., Majumdar, D. and Sen, A. (2015). Incentive-compatible voting rules with positively correlated beliefs, *Theoretical Economics* **10**(3): 867–885.

Brady v. Maryland (1963). U.S. Supreme Court.

Cella, M. (2008). Informed Principal with Correlation, *Games and Economic Behavior* **64**(2): 433–456.

Crémer, J. and McLean, R. P. (1988). Full Extraction of the Surplus in Bayesian and Dominant Strategy Auctions, *Econometrica* **56**(6): 1247–1257.

Demougin, D. M. and Garvie, D. A. (1991). Contractual Design with Correlated Information under Limited Liability, *The RAND Journal of Economics* pp. 477–489.

Erlanson, A. and Kleiner, A. (2020). Costly Verification in Collective Decisions, *Theoretical Economics* **15**(3): 923–954.

Franz II (1803). Gesetzbuch über Verbrechen und Schwere Polizey-Uebertretungen vom 3. September 1803, *Zweyter Theil, Von den schweren Polizey-Uebertretungen* pp. 115–125.

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

Gershkov, A., Goeree, J. K., Kushnir, A., Moldovanu, B. and Shi, X. (2013). On the Equivalence of Bayesian and Dominant Strategy Implementation, *Econometrica* **81**(1): 197–220.

Gibbard, A. (1973). Manipulation of voting schemes: a general result, *Econometrica* **41**(4): 587–601.

Glazer, J. and Rubinstein, A. (2004). On Optimal Rules of Persuasion, *Econometrica* **72**(6): 1715–1736.

Halac, M. and Yared, P. (2020). Commitment vs. Flexibility with Costly Verification, *Journal of Political Economy* (forthcoming).

Hart, S., Kremer, I. and Perry, M. (2017). Evidence Games: Truth and Commitment, *The American Economic Review* **107**(3): 690–713.

Johnson, S., Pratt, J. W. and Zeckhauser, R. J. (1990). Efficiency despite mutually payoff-relevant private information: The finite case, *Econometrica* pp. 873–900.

Kittler, W. (2003). Heimlichkeit und Schriftlichkeit: Das österreichische Strafprozessrecht in Franz Kafkas Roman Der Proceş, *The Germanic Review: Literature, Culture, Theory* **78**(3): 194–222.

Manelli, A. M. and Vincent, D. R. (2010). Bayesian and Dominant-Strategy Implementation in the Independent Private-Values Model, *Econometrica* **78**(6): 1905–1938.

Maskin, E. and Tirole, J. (1990). The Principal-Agent Relationship with an Informed Principal: The case of private values, *Econometrica* pp. 379–409.

McAfee, R. P. and Reny, P. J. (1992). Correlated Information and Mechanism Design, *Econometrica* pp. 395–421.

Milgrom, P. R. and Weber, R. J. (1982). A Theory of Auctions and Competitive Bidding, *Econometrica* pp. 1089–1122.

Myerson, R. B. (1981). Optimal Auction Design, *Mathematics of Operations Research* **6**(1): 58–73.

Myerson, R. B. (1983). Mechanism Design by an Informed Principal, *Econometrica* pp. 1767–1797.

Mylovanov, T. and Zapechelnyuk, A. (2017). Optimal Allocation with Ex Post Verification and Limited Penalties, *American Economic Review* **107**(9): 2666–94.

Neeman, Z. (2004). The relevance of private information in mechanism design, *Journal of Economic Theory* **117**(1): 55–77.

Riordan, M. H. and Sappington, D. E. (1988). Optimal Contracts with Public ex post Information, *Journal of Economic Theory* **45**(1): 189–199.

Satterthwaite, M. A. (1975). Strategy-proofness and Arrow's conditions: Existence and correspondence theorems for voting procedures and social welfare functions, *Journal of Economic Theory* **10**(2): 187–217.

Severinov, S. (2008). An Efficient Solution to the Informed Principal Problem, *Journal of Economic Theory* **141**(1): 114–133.

Townsend, R. M. (1979). Optimal Contracts and Competitive Markets with Costly State Verification, *Journal of Economic Theory* **21**(2): 265–293.

3
Inspecting Experimentation

3.1 Introduction

Uncertainty and progressive learning are inherent features of innovative ventures. They intensify the incentive problem that arises when fund provision and control of the venture are separated. The experimenter can misappropriate the provided funds for private consumption or inefficient investments that yield private benefits. One tool to handle this dynamic moral hazard problem is the design of bonuses that reward the agent for positive outcomes to align his interests with those of the principal. This paper introduces *inspections* as an additional tool. In practice, investors spend considerable time and effort to make the resources put into experimentation public.

> Venture capitalists concentrate investments in early stage companies and high technology industries where informational asymmetries are significant and monitoring is valuable. (Gompers and Lerner, 2004, pp. 132-133)

Examples beyond venture capital financing include the allocation of funds to universities by the government and the distribution of the R&D budget to different projects within an organisation. Processing information on current and past fund allocation entails an important fixed cost component. When inspectors travel to the agent's facilities for interviews, verify the state of prototypes, or create progress reports, the required resources vary only little in the time-frame and amount of funds to be accounted for.

> Because monitoring is costly and cannot be performed continuously, the venture capitalist will periodically check the project's status and preserve the option to abandon. (Gompers and Lerner, 2004, p. 139)

The aim of this paper is to analyse optimal experimentation contracts, combining inspections and payments to avert fund diversion at minimal cost. Should inspections be random or deterministic and what is the optimal timing? What is the optimal payment schedule and contract length and, in particular, how are these affected by inspections?

Analysis. To answer the above questions, I study the experimentation relationship in a two-armed bandit model.[1] A principal provides funds to an agent to experiment on a project of uncertain quality. In each period, the agent privately decides whether to experiment on the project or divert the funds provided by the principal for his private benefit. If the agent experiments, the good project yields a publicly observable success with some probability. The agent's experimentation choices are not observable by the principal and need to be incentivised through payments, potentially contingent on success. The agent is protected by limited liability so that all payments have to be non-negative. The novel feature is that the principal can perform inspections which allow her to detect fund diversion with some probability that is increasing in the amount of funds the agent has appropriated. The principal can commit to a contract, and the projects success and inspection outcomes are contractible.

To build intuition, I first characterise the optimal bonus stream and contract duration for the benchmark case without inspections. It is well known from the literature that the optimal contract requires the agent to continuously experiment in every period until success occurs or some finite deadline is reached. The players gradually become more pessimistic about the project's quality if the agent experiments and produces no success. This implies that, after diverting funds and not experimenting for some time, the agent is more optimistic than his (on-path) self who experimented without success. Hence, the most profitable deviations – which determine the bonus payments – are to divert funds for a short time and experiment afterwards. The bonus payments leave rents to the agent that consist of two components. First, since experimentation may lead to a success and thereby end the relationship, doing so limits the agent's option to divert funds in the future. For this, he is compensated with a dynamic rent. Second, the more optimistic belief about the project's quality that would result from fund-diversion requires an information rent. The total rent increases exponentially in the remaining length of the contract. Therefore, the principal commits to ending the relationship inefficiently early. If players could renegotiate at the end of the contract, they would agree to experiment for longer.

[1] Thereby the paper builds on Bergemann and Hege (1998), Keller et al. (2005), and Hörner and Samuelson (2013), among others. For a survey, see Bergemann and Välimäki (2008).

With inspections, optimal contracts still require the agent to experiment for the entire contract duration. If an inspection detects otherwise, the relationship is terminated since this provides the harshest punishment the principal can apply. The optimal payment scheme after inspection is given by the non-inspection benchmark as the conditions to deter fund diversion are equivalent. Prior to inspection, the optimal payment rule depends crucially on the precision of the inspection technology. An inspection that results in detection gives a conclusive signal of fund diversion, the arrival of a success gives a conclusive signal of experimentation; the lack of either detection or success provides a non-conclusive signal. If the precision of the signal resulting from the lack of success exceeds the precision of the signal resulting from lack of detection, bonus payments are necessary at all times for incentive provision. However, the dynamic rent is reduced as inspections make the option of future fund diversion less likely. In the other case, if the precision of inspections exceeds a threshold, it is optimal to shift all payments to the future and pay bonuses only for successes after inspections. Dynamic rents are decreased even further in this case as the principal leverages the threat of losing post-inspection rents to incentivise pre-inspection experimentation.

Under this *re-use* of rents, the optimal contract duration is strictly larger than in the benchmark without inspections. This effect can be strong enough to make the principal commit to a longer contract duration than would be statically optimal. That is, if the principal had not committed to the termination deadline, starting from some time prior to the deadline, she would like to renegotiate and stop immediately. The agent would reject this so the contract is renegotiation proof. In any case, the contract length never exceeds the first best duration of experimentation.

The optimal timing of inspection with precise inspections is entirely deterministic. Randomisation reduces the threat that an inspection poses to the fund-diverting agent for the following reason. With experimentation, conditioning future times on the event that no success arrived previously leads both players to effectively discount the future at a higher rate. As fund diversion rules out success, a deviating agent has a lower effective discount rate. Lower discount rates imply higher risk aversion over time-lotteries (DeJarnette et al., 2020) so that a discounted-mean-preserving contraction of inspection times intensifies the expected threat of an inspection for the diverting agent.

The remainder of the paper is organised as follows. The next section presents the model, defines the class of feasible contracts, and describes the resulting payoffs. Optimal contracts are discussed in Section 3.3 for the case without inspections in Section 3.4 for the case with inspections. Concluding remarks are presented in Section 3.5. All proofs are contained in the Appendix 3.A. Apart from the papers mentioned here, additional literature on dynamic inspections can be found in 4.1.

3.2 Model

Environment. There is a principal (she) who contracts with an agent (he) to experiment on a project in continuous time $t \in [0, \infty)$. At every instant, the principal chooses whether to provide funds of $\varphi > 0$ which are necessary for experimentation. The agent has no funds of his own and is protected by limited liability so that all payments from principal to agent have to be non-negative. When given the necessary funds, the agent decides whether to invest them to experiment on the project ($a_t = 1$) or to divert the funds provided for that instant ($a_t = 0$).

The project is either of good or bad quality, denoted by $\omega \in \{0, 1\}$, where $\omega = 1$ encodes the good state. Both players share initial prior belief $\mu_0 \in (0, 1)$ that the project is good. At every instant, the project yields a success with Poisson arrival rate $\omega\lambda a_t$. That is, a bad project yields success with probability 0, independent of the agent's experimentation choice. A good project's success rate is λ if the agent experiments, and 0 if he does not experiment. Success is publicly observable. The first success ends the game and gives the principal utility normalised to 1.[2] The agent does not intrinsically care about success. When he shirks and diverts the provided funds, he obtains an instant flow payoff of φ.[3]

The agent's experimentation choice is hidden. The principal has two instruments to incentivise it. First, she can offer monetary payments in the form of fixed wages as well as bonuses that are paid conditional on the arrival of a success. Second, the principal can schedule inspections. At any inspection time $t_I \geq 0$ she observes detection state $\theta_{t_I} \in \{0, 1\}$ where $\theta_0 = 0$, θ_t transitions from 0 to 1 at exponential rate $\delta(1 - a_t)$, and the state $\theta_t = 1$ is absorbing. That is, the *detection rate* $\delta \geq 0$ captures the precision of the inspection technology. If $\delta = 0$, $\theta_t = 0$ forever so that inspections are uninformative. In the limit as $\delta \to \infty$, any positive amount of funds the agent diverts is detected at the next inspection with certainty. The interpretation of this technology is that, whenever the agent diverts funds, he risks 'leaving a paper-trail' with probability proportional to the amount of funds he diverted. If the trail is detectable ($\theta_t = 1$) it is found at the next inspection. The state θ_t is observed neither by the principal nor by the agent prior to inspections. At the time of each inspection, the principal pays fixed cost $\kappa > 0$.

Both players are risk neutral expected-utility maximisers. If one player decides not to participate in the relationship at the beginning, both receive utility 0.

[2] One can think of 1 as the discounted expected value of continuing a project known to be good to cover the case of multiple valuable successes.

[3] The analysis is easily adapted to the case where the agent receives utility $\psi \neq \varphi$ from diverting the funds. Specifically, $\psi < \varphi$ captures the case where only part of the funds can be appropriated by the agent.

Contracts. At the beginning of the relationship the principal commits to a contract to maximise her expected utility. Two prior observations are useful to simplify the class of admissible contracts. First, any optimal contract requires the agent to experiment at every instant until either a success occurs or some deadline is reached. Second, in the off-equilibrium event that inspection reveals fund-diversion, the contract specifies the harshest possible punishment by terminating the relationship immediately.[4] Hence, without loss of generality, the principal offers a contract $\mathcal{C} = (T, N^I, b, w)$ consisting of the following components:

i) A *deadline* $T \in [0, \infty]$ such that the agent is expected to work at every instant $t \in [0, T]$ if no success has arrived previously and no fund diversion was detected.

ii) An *inspection process* $N^I = \left(N_t^I\right)_{t=0}^{T} \in \mathbb{N}_0^{[0,T]}$ which is non-decreasing, right-continuous and has $dN_t^I = 1$ if an inspection occurs at t and $dN_t^I = 0$ otherwise. N_t^I and $dN_t^I \theta_t$ are public.

iii) A *bonus* schedule $b = (b_t)_{t=0}^{T} \in \mathbb{R}_+^{[0,T]}$. Bonus b_t determines the payment the agent receives immediately after success in period t. Accordingly, the principal receives $1 - b_t$.

iv) A *wage* process $w = (w_t)_{t=0}^{T}$ which is non-decreasing with $dw_t \equiv w_t - \lim_{s \uparrow t} w_s$.[5]

The game ends with the arrival of a success or at time T. Deadline and inspections may be random. All events are to be understood as *conditional* on no success and no detection having occurred previously as these events end the interaction. Part (iii) states that, in case of success, the promised bonus is paid to the agent immediately. In particular, the principal does not profit from performing any inspection after success occurred. An equivalent interpretation of b_t is that of the agent's share of the venture at time t.

Agent. Given contract $\mathcal{C}$, the agent chooses an action plan $a = (a_t)_{t=0}^{T}$ with the restriction that a be absolutely continuous with respect to the Lebesgue measure on $[0, T]$ and measurable with respect to the natural filtration induced by inspection process N^I. As above, a is to be understood as conditional on continuing the relationship, that is, conditional on no detection or success. To economise on notation, a denotes both the agent's strategy and the resulting path realisation. In the case of deterministic inspections, they are identical.

[4] The first observation follows from the fact that no information is created if $a_t = 0$ on some positive interval so that the principal can close any gaps without altering the agent's incentives (see Lemma A.3 in Green and Taylor, 2016, for a formal argument). The second observation follows from the agent's limited liability and the fact that shirking does not occur on the equilibrium path.

[5] The optimal contract will feature no fixed wages. Including them in the contract specification facilitates the exposition.

Define detection (stopping) time $\chi = \inf\{t \in [0, T] : dN_t\theta_t = 1\}$ with the convention that $\chi = T$ if $dN_t\theta_t = 0$ always. The distribution of χ depends on the agent's strategy. The agent's expected utility resulting from contract $\mathcal{C}$ and strategy a is

$$U(\mathcal{C}, a) =$$
$$\mathbb{E}^a_\chi \left[\int_0^\chi \left[\mu_0 e^{-\lambda \int_0^t a_s ds} \lambda a_t b_t dt + \left(\mu_0 e^{-\lambda \int_0^t a_s ds} + (1-\mu_0) \right) (dw_t + (1-a_t)\varphi dt) \right] \right],$$

where the sub- and superscript are included to highlight the dependence of χ on a. The first term in the square brackets captures the utility from the bonuses. If the state is good, time t is reached with probability $e^{-\lambda \int_0^t a_s ds}$. A success arrives at rate λa_t. The second term captures the agent's payoff from wages and fund diversion which accrue independently of the state as long as no success occurred previously.

Principal. Assume that the contract $\mathcal{C} = \left(T, N^I, b, w\right)$ is such that it is optimal for the agent to experiment throughout. Then the principal's expected discounted payoff from contract $\mathcal{C}$ is

$$V(\mathcal{C}) = \int_0^T \left[\mu_0 e^{-\lambda t} \lambda (1-b_t) dt + \left(\mu_0 e^{-\lambda t} + 1 - \mu_0 \right) \left(-dw_t - \varphi dt - \kappa dN_t^I \right) \right].$$

Conditional on the project being good, every period t is reached with probability $e^{-\lambda t}$, success arrives at rate λ, and the principal receives expected payoff $(1 - b_t)$. Prior to a success, the principal pays the stream of wages, fund provision, and inspection cost κ whenever $dN_t^I = 1$.

The principal's problem can be stated as

$$\max_{\mathcal{C}} \; V(\mathcal{C}) \text{ s.t.}$$
$$\mathbf{1} \in \arg\max_a U(\mathcal{C}, a). \qquad (IC)$$

Here, $\mathbf{1}$ denotes the process a with $a_t = 1$ everywhere. Likewise, $\mathbf{0}$ will be used below to denote processes that are constantly 0 in the case of w and N^I. The agent can always secure utility of at least 0 by choosing $a_t = 0$ for all t so his participation constraint is automatically fulfilled in any incentive compatible contract.

3.3 Optimal Contract without Inspections

This section presents the solution to the principal's problem in the no-inspection benchmark in which she is forced to choose $N_t^I = 0$ for all t. The solution to this problem is known in the literature. With slight variations in the model setup, it can be found in Bergemann and Hege (1998), Hörner and Samuelson (2013), Green and Taylor (2016), and Halac et al. (2016), among others. Therefore, I anticipate that the optimal contract features no fixed wages and illustrate the optimal bonuses and deadlines with intuitive arguments here. Proposition 4 in Appendix 3.A.1 contains the formal result. Its proof relies on techniques from optimal impulse control; see Arutyunov et al. (2018).

Bonuses. Consider first the optimal bonus payments for fixed deadline T. Let μ_t denote the posterior belief if, despite continual experimentation, no success has arrived until time t. By Bayes' rule,

$$\mu_t = \frac{\mu_0 e^{-\lambda t}}{1 - \mu_0 + \mu_0 e^{-\lambda t}}.$$

The analysis below will make repeated use of the identities $\frac{1-\mu_0}{1-\mu_t} = \mu_0 e^{-\lambda t} + 1 - \mu_0$ and $\frac{1-\mu_0}{1-\mu_t}\mu_t = \mu_0 e^{-\lambda t}$. The principal's problem for optimal bonuses reads

$$\max_b \left\{ \int_0^T \frac{1-\mu_0}{1-\mu_t} \left(\mu_t \lambda (1 - b_t) - \varphi\right) dt \right\} \quad \text{s.t.}$$
$$\mathbf{1} \in \arg\max_a U\left((T, \mathbf{0}, b, \mathbf{0}), a\right).$$

The bonuses necessary to incentivise experimentation in every period can be constructed recursively. Which potential deviations determine the relevant incentive constraints? If the bonuses are such that the agent who has experimented until period t is willing to experiment from $t + dt$ until T, then after diverting the funds during $[t, t + dt)$, he has even more incentive to experiment. The reason is that his belief about project quality is more optimistic after fund diversion than after experimenting without success. Define by z_t the agent's expected discounted utility from honouring the contract *conditional* on the project being good:

$$z_t = \int_t^T e^{-\lambda(s-t)} \lambda b_s ds. \tag{3.1}$$

At belief μ_t, the agent's expected discounted utility is $\mu_t z_t$. The resulting incentive constraint to avert fund diversion during $[t, t + dt)$ is

$$\mu_t z_t \geq \varphi dt + \mu_t z_{t+dt}. \tag{3.2}$$

In the optimal contract, this constraint is binding for all t. Applying a Taylor-expansion and discarding terms that vanish faster than dt, one obtains the condition $0 = \frac{\varphi}{\mu_t} + z_t'$. Inserting μ_t and the boundary condition $z_T = 0$, we get the *benchmark* *continuation utility* at time t given deadline T:

$$z_t^B = \varphi \left(T - t + \frac{1-\mu_0}{\mu_0} \frac{e^{\lambda T} - e^{\lambda t}}{\lambda} \right). \tag{3.3}$$

The deadline is omitted from the definition of z_t^B unless the dependence shall be highlighted explicitly. To obtain the corresponding bonuses, take the derivative in (3.1) to get $z_t' = -\lambda b_t + \lambda z_t$. Inserting z_t^B, the optimal *benchmark bonus* schedule is

$$b_t^B = \frac{\varphi}{\mu_t \lambda} + \varphi(T-t) + \varphi \frac{1-\mu_t}{\mu_t} \frac{e^{\lambda(T-t)} - 1}{\lambda}.$$

Bonuses prior to the deadline have to compensate not only for the forgone shirking payoff with amount $\frac{\varphi}{\mu_t \lambda}$ but also offer a dynamic rent $\varphi(T-t)$ to compensate the agent for the fact that experimentation may end the relationship, depriving him of the option to divert funds until T. However, if the agent shirked at time t, he would strictly prefer to experiment so that his forgone utility in case of success is larger than $\varphi(T-t)$. The last term captures this additional information rent.

Contract Duration. Before determining the optimal contract duration without inspections, suppose first that the principal acts myopically and consider the last instant of a contract. At $t = T$, the bonus is $b_T^B = \frac{\varphi}{\mu_T \lambda}$. Ignoring previous periods, the principal is willing to provide funds and offer bonus b_T for experimentation if

$$-\varphi dt + \mu_T \lambda dt \, (1 - b_T) \geq 0.$$

Inserting the bonus from above, this is equivalent to $\mu_T \lambda \geq 2\varphi$. Abstracting from dynamic rents, the principal is willing to finance experimentation whenever the expected proceeds from the project ($\mu_T \lambda$) exceed the cost of funding (φ) plus the cost of incentives (φ). Define the *static deadline* as[6]

$$T^{\text{static}} \equiv \sup\{T' \geq 0 : \mu'_T \lambda \geq 2\varphi\}.$$

We will compare this static deadline, which will generally not be optimal, to the optimal deadlines *without* and *with* inspections. For the remainder of the analysis, assume that the principal wants to experiment at least initially, i.e. $\mu_0 \lambda - 2\varphi > 0$.

When choosing the optimal deadline, the principal internalises that each instant of experimentation raises the bonuses at previous times due to the increasing rents. Given the agent's payoff as a function of the deadline in (3.3), the optimal deadline

[6] The myopic problem is considered as well in Green and Taylor (2016). In the current paper, the condition features 2φ due to the fact that fund diversion yields utility of the same amount. If the funds delivered utility of ψ, the condition would be $\mu_T \lambda \geq \varphi + \psi$, as in Green and Taylor (2016).

can be derived by solving the following problem:

$$\max_{T\geq 0}\left\{\int_0^T \frac{1-\mu_0}{1-\mu_t}\left(\mu_t\lambda-\varphi\right)dt-\mu_0 z_0^B(T)\right\}.$$

It is easily verified that the objective is concave in T and strictly positive whenever $\mu_0\lambda > 2\varphi$. Using $z'(t) = -\frac{\varphi}{\mu_t}$, the problem can be rewritten as

$$\max_{T\geq 0}\left\{\int_0^T \frac{1-\mu_0}{1-\mu_t}\left(\mu_t\lambda-\varphi-\varphi e^{\lambda t}\right)dt\right\}.$$

The bracketed term in the integral is strictly decreasing in t, so that the optimal contract continues as long as it is above 0. Define the optimal *benchmark deadline* by T^B. It is determined by the solution to

$$\lambda\mu_{T^B} = \varphi + \varphi e^{\lambda T^B}. \tag{3.4}$$

The expected proceeds ($\mu_T\lambda$) have to compensate the principal for the cost of funding (φ) and the incentive cost ($\varphi e^{\lambda T}$) which now includes the dynamic rent component and is higher that in the myopic consideration. As the belief is decreasing over time, it follows that $T^B < T^{\text{static}}$. Without inspections, the principal commits to ending the relationship earlier than she would myopically. This implies that the contract is not renegotiation proof: At time T^B, the principal would prefer to have the agent experiment further and pay the necessary continuation utility. For this reason, replacement of the entrepreneur has been suggested to improve efficiency in experimentation settings (Bergemann and Hege, 1998). The next section shows that this may be reversed with inspections, the principal sometimes commits to experimenting longer than myopically optimal. If that is the case, the contract becomes renegotiation proof as the agent would not agree to forgo his continuation value by earlier termination and, given the principal has to provide a positive payoff, she prefers to do so through a contract rather than simply handing out a lump sum payment.

3.4 Optimal Contract with Inspections

I focus on the case in which the principal can schedule at most one inspection. Given the results in the previous section, denote by $V^0(\mu_t, z)$ the principal's value function without inspections at initial belief μ_t and agent's continuation utility z. The formal expression is given in (3.A.1) in the appendix. The principal's problem in case of a single inspection can be written as

$$V^1(\mu_0, u_0) =$$
$$\sup_{t_1,F,b,z} \left\{ \int_0^{t_1} \frac{1-\mu_0}{1-\mu_t} \Big[(1-F_t)\left((\mu_t\lambda(1-b_t)-\varphi)dt - dw_t\right) + \Big(V^0(\mu_t, z_t) - \kappa\Big)\, dF_t\Big] \right\},$$
$$\text{s.t. } (PK)\colon \mu_0 u_0 = \int_0^{t_1} \frac{1-\mu_0}{1-\mu_t} \left[(1-F_t)\left(\mu_t\lambda b_t dt + dw_t\right) + \mu_t z_t dF_t\right],$$
$$(IC).$$

Here, t_1 is the end date of the first (pre-inspection) stage; without loss we can assume that $t_1 \leq T^{FB}$ which is defined by $\mu_{T^{FB}}\lambda = \varphi$.[7] F is the cdf of the inspection-time distribution $\in \Delta([0, t_1])$. Processes b_t and z_t denote the bonuses and continuation utility in case of success or inspection at t.

As in the previous section, the results on the optimal contract components are established consecutively. I first consider payments, then contract duration, and finally the inspection policy.

Payments. The result on all contract components will depend significantly on the precision of the inspection technology measured by δ. For payments, we have the following result:

Proposition 1. *The payments in the optimal contract with at most one inspection are as follows:*

- *There are no fixed wages, $w_t^* = 0$ for all t.*
- *If $\delta \leq \lambda$, the optimal contract features positive bonuses throughout: $b_t^* > 0$ for all t.*
- *There exists $\bar{\delta} \in (\lambda, \infty)$ such that, if $\delta \geq \bar{\delta}$ and the optimal contract includes inspections, then no bonuses are paid prior to inspection.*

The formal proof of this and all following results are relegated to the appendix and derived by means of the optimal-control techniques in Arutyunov et al. (2018).

The intuition for the result is as follows. Consider first the case $\delta \leq \lambda$. If the agent diverts funds from time 0 until an inspection at time t, he gets the continuation utility z_t with probability $e^{-\delta t}$, the probability that he is not detected. If the agent experiments, he gets continuation utility z_t only if no success occurs previously, i.e. with probability $e^{-\lambda t}$. If $\lambda \geq \delta$, the second probability does not exceed the first and

[7] After T^{FB}, the principal's expected return from experimentation is negative. She is better off by cutting off the remaining part of the contract. If the original contract includes an inspection after T^{FB}, the inspection can be moved to T^{FB} with lower probability replacing the agent's continuation utility from the remaining contract with a cash payment.

termination before t is triggered more likely by experimentation and success than by diversion and detection. If no bonuses were paid, the agent would prefer to reap the benefit of fund diversion and obtain the continuation utility at t with higher probability.

Next, let δ be large. On path, principal and agent have the same belief so that payments can be shifted into the future in a ratio that maintains the principal and the (non-deviating) agent indifferent. If the detection rate is high enough, this makes fund diversion prior to inspection less attractive. While it increases the probability of reaching the inspection date without success and at a more optimistic belief, the agent risks being detected. When the detection rate δ exceeds success rate λ by a sufficient amount, the second effect dominates. In this case, the rents necessary for post-inspection incentives have an additional reward for the principal. The threat of loosing the promised continuation payoff at inspection prevents fund diversion prior to inspection without additional payments. The rents from the post-inspection stage, which give the agent strictly positive continuation payoff, are re-used to provide incentives also at the pre-inspection stage.

Contract Duration. Due to the re-use of rents, the principal's tendency to commit to inefficient early termination is dampened with inspections, as the following result shows.

Proposition 2. *Suppose an inspection is performed at time t^*. Let $T^* - t^*$ be the remaining contract duration. If $\delta > \lambda$, then $T^* - t^* > T^B(\mu_{t^*})$, the continuation contract lasts longer than the optimal no-inspection benchmark starting with prior belief μ_{t^*}.*

To compare the two deadlines, suppose we start at time 0 but with prior belief μ_{t^*} and choose the optimal deadline T^B without inspection. Benchmark deadline T^B was determined in (3.4) by the condition $\mu_{T^B}\lambda - \varphi - \varphi e^{\lambda T^B} = 0$. For the case with inspection, the proof in the appendix reveals that there is a multiplier $\psi(t) > 0$ associated with the agent's incentive constraint such that, when inspection occurs at time t^*, experimentation continues optimally until the time T^* determined by

$$\mu_{T^*}\lambda - \varphi - \varphi e^{\lambda(T^*-t^*)} + \varphi e^{\lambda(T^*-t^*)}\psi(t)(1 - e^{-(\delta-\lambda)t^*}) = 0.$$

While the first three terms are the same as in the optimality condition in the no-inspection benchmark when starting at t^*, there is an additional term, which is positive when $\delta > \lambda$ so that T^* has to be strictly later than T^B. Recall that we defined T^{static} as the solution to $\mu_t - 2\varphi = 0$. Hence, when $-\varphi e^{\lambda(T^*-t^*)} + \varphi e^{\lambda(T^*-t^*)}\psi(t)(1 - e^{-(\delta-\lambda)t^*}) > -\varphi$, the principal values the rents created in the continuation contract so much that she commits to experimenting longer than myopically optimal.

Inspection Policy. The previous two results reveal how payments and deadline are chosen optimally. What is the optimal distribution over inspection times? Perhaps surprisingly, the optimal inspection time is deterministic when inspections are sufficiently precise.

Proposition 3. *If the inspection cost κ exceeds $\bar{\kappa}$, the benchmark contract without inspections is optimal. Otherwise, there is a principal-optimal contract with inspection.*

- *The support of the inspection distribution F^* is finite.*
- *If $\delta > \bar{\delta}$, then there is a deterministic inspection time t^* and deadline T^* such that there are no payments prior to t^*, and the continuation contract from t^* onward is given by the benchmark contract starting at belief μ_{t^*} with remaining duration $T^* - t^*$.*

The reason why randomisation does not help in incentive provision lies in the time-risk preferences induced by the succession of events in an experimentation setting and how the agent's choices affect their arrival rates. Consider a random inspection time $\tilde{t}$. When the agent experiments, payoffs at time $\tilde{t}$ (conditional on no prior success) are effectively discounted at rate λ. The expected discount factor is $\mathbb{E}\left[e^{-\lambda\tilde{t}}\right]$. When the agent diverts funds, $\tilde{t}$ is reached for certain without prior success. The expected time is $\mathbb{E}\left[\tilde{t}\right]$. Consider replacing $\tilde{t}$ with deterministic time t such that $e^{-\lambda t} = \mathbb{E}\left[e^{-\lambda\tilde{t}}\right]$, so the expected discount factor remains equal for the principal and the agent who experiments. The convexity of the exponential function and Jensen's inequality imply that $t \leq \mathbb{E}\left[\tilde{t}\right]$, with strict inequality if $\tilde{t}$ is non-degenerate. The funds the agent can divert until inspection decrease from $\varphi\mathbb{E}\left[\tilde{t}\right]$ to φt while the expectations for principal and on-path agent remain equal. Thus, replacing $\tilde{t}$ by the deterministic inspection time t reduces the agent's incentives to divert funds while leaving the expected discounted experimentation payoff unchanged.[8] DeJarnette et al. (2020) show that, under exponential discounting, impatience induces risk-seeking over time lotteries.[9] Discounting increases the expected impact of payoffs that are timed randomly. In the present setting, an inspection has a negative impact on the agent's payoff as it terminates his fund-diversion opportunity. when he diverts funds, the principal's effective discount rate is higher than the agent's. Hence, when its timing is deterministic, the expected impact of the inspection is relatively larger on the agent.

[8] Note that this comparison is not driven by the absence of standard discounting. If we want to capture the players' impatience with discount rate r, the above change from random to deterministic timing has the same effect, where the linear term corresponding to the perspective of the diverting agent is replaced by e^{-rt} and the term corresponding to on-path behaviour is replaced by $e^{-(r+\lambda)t}$.

[9] A time lottery in DeJarnette et al. (2020) is a fixed (positive) lump-sum payment realised at a random time.

3.5 Concluding Remarks

This paper incorporates the possibility of detecting fund-diversion in a two-armed bandit model of strategic experimentation. I analyse the effect of the principal's inspection ability on optimal contracts. I find that the optimal contract depends crucially on the quality of the inspection technology. With low inspection precision, bonuses are required for incentives at all times. With precise inspections, early periods are incentivised solely through inspections, re-using the rents resulting from the later payments that are required for incentives after inspection. In this case, the optimal inspection timing is deterministic. The next chapter of this thesis reveals that randomisation can be valuable for incentive provision in non-experimentation settings and and shows that the discounting-induced risk-preferences of a diverting agent can go in the other direction. Possible further steps within the present study include the characterisation of optimal contracts for low and intermediate inspection precision and the principal's freedom to choose the total number of inspections.

3.A Appendix

This appendix contains the formal proofs not included in the main text. I derive necessary condition applying the Maximum Principal in Arutyunov et al. (2018) which allows for continuous and impulse controls. Existence of a solution is ensured by Theorem 7.1 therein.

3.A.1 Proofs without Inspections

The following result characterises the optimal contract underlying Section 3.3.

Proposition 4. *Without inspections, given initial belief μ and promised utility z, the principal's value function is*

$$V^0(\mu,z) = -\mu z + \mu\frac{1-e^{-\lambda T(\mu,z)}}{\lambda}(\lambda-\varphi) - (1-\mu)\varphi T(\mu,z), \tag{3.A.1}$$

with $T(\mu,z) = \min\left\{T^{FB}; \frac{z}{\varphi} + \frac{1-\mu}{\mu\lambda} - \frac{1}{\lambda}W\left(\frac{1-\mu}{\mu}e^{\frac{1-\mu}{\mu}+z\lambda/\varphi}\right)\right\}$. *Here, $W(x)$ denotes the product logarithm which is implicitly defined as the value w s.t. $we^w = x$.*

Proof. To match the (Meyer) formulation of the control problem in Arutyunov et al. (2018), define the additional state O_t where O stands for the principal's objective. The maximum problem can be written as

$$\min_{T,b,w}\{-O(T)+w_0\}$$

s.t.

$$dO_t = \frac{1-\mu_0}{1-\mu_t}\left((\mu_t\lambda(1-b_t)-\varphi)dt - dw_t\right) \qquad O_0 = 0$$

$$dz_t = \lambda(z_t-b_t)dt - \frac{dw_t}{\mu_t} \qquad z_0 = z0, z_T \geq 0$$

$$\lambda b_t \geq \frac{\varphi}{\mu_t} + \lambda z_t.$$

Let ψ^X denote the co-state associated with state $X \in \{O, z\}$. The Hamiltonian associated with this problem is given by

$$H(t) = \psi_t^O\frac{1-\mu_0}{1-\mu_t}(\mu_t\lambda(1-b_t)-\varphi) + \psi_t^z\lambda(z_t-b_t).$$

The impulse function, Q, associated with the control dw_t is given by

$$Q(t) = -\psi_t^O\frac{1-\mu_0}{1-\mu_t} - \psi_t^z\frac{1}{\mu_t}.$$

Optimality demands $Q(t) \leq 0$ for all $t \in [0,T]$ with equality whenever $dw_t > 0$. The co-state ψ_t^O is constant as O_t is absent from both H and Q and we have $\psi_t^O = \psi^O \in \{0,1\}$. The co-state ψ_t^z satisfies $\psi_T^z \geq 0$ and $d\psi_t^z = -\psi_t^z\lambda dt$, so that $\psi_t = e^{\lambda(T-t)}\psi_T^z$. This implies that

$$\begin{aligned}Q(t) &= -\psi^0(\mu_0e^{-\lambda t}+1-\mu_0) - e^{\lambda(T-t)}\psi_T^z(1+\frac{1-\mu_0}{\mu_0}e^{\lambda t})\\ &= -\left(\mu_0e^{-\lambda t}+1-\mu_0\right)\left(\psi^0+\frac{\psi_0^z}{\mu_0}\right)\end{aligned}$$

By optimality conditions (6.17-20) in Arutyunov et al. (2018, pp. 138-139), $\psi^O + |\psi_0^z| > 0$ and $\psi^O, \psi_T^z \geq 0$. Therefore, we have $Q(t) < 0$ for all t and $dw_t = 0$ for all t.

This gives a simplification of the problem, as only an initial wage w_0 may be paid at the beginning of the contract. Further, constraint $\lambda b_t \geq \frac{\varphi}{\mu_t} + \lambda z_t$ has to bind everywhere as otherwise we need $\frac{\partial H}{\partial b} = 0$ which implies $\psi^O + \psi_0^z = 0$, a contradiction.

The problem is reduced to choosing the optimal initial wage w_0 and deadline T to maximise

$$\int_0^T \frac{1-\mu_t}{1-\mu_0} (\mu_t \lambda - \varphi)\, dt,$$

subject to the constraint $\mu_0 z_0 = w_0 + \varphi \int_0^T \frac{1-\mu_t}{1-\mu_0} e^{\lambda t} \lambda dt$. It is optimal to choose $w_0 = 0$ if $\mu_T \lambda - \varphi > 0$ for $T : \varphi \int_0^T \frac{1-\mu_t}{1-\mu_0} e^{\lambda t} dt = \mu_0 z_0$. Otherwise, it is optimal to choose $T = T^{FB} \equiv T' : \mu_{T'} \lambda - \varphi = 0$ and $w_0 = \mu_0 z_0 - \varphi \int_0^{T^{FB}} \frac{1-\mu_t}{1-\mu_0} e^{\lambda t} dt$. □

3.A.2 Proofs with Inspection

To match the (Meyer) formulation of the control problem in Arutyunov et al. (2018), define the additional state O_t where O stands for the principal's objective. Further, I do not include wages as controls and verify in the proof of Proposition 1 that this is optimal. The control problem reads

$$\begin{aligned}
&\min_{t_1, \nu, z} \{-O(t_1)\} \\
&dO_t = (1-F_t)\frac{1-\mu_0}{1-\mu_t}\left[(\mu_t \lambda(1-b_t) - \varphi)dt\right] + \frac{1-\mu_0}{1-\mu_t}\left(-\kappa + V^0(\mu_t, z_t)\right) d\nu_t && O_0 = 0 \\
&dB_t = (B_t - (1-F_t)b_t)\,\lambda dt && B_0 \geq 0, B_{t_1} = 0 \\
&dZ_t = Z_t \lambda dt - z_t d\nu_t && Z_0 \geq 0, Z_{t_1} = 0 \\
&dF_t = d\nu_t && F_0 = 0, F_{t_1} \leq 1 \\
&dG_t = (1-F_t)\left(\mu_0 e^{-\lambda t} \lambda b_t - \varphi\right) dt + \mu_0 \left(e^{-\lambda t} - e^{-\delta t}\right) z_t d\nu_t && G_0 = 0, G_{t_1} \geq 0 \\
&\ell(b_t, F_t, B_t, Z_t, t) = -\lambda b_t (1-F_t) + \frac{\varphi}{\mu_t}(1-F_t) + \lambda B_t - (\delta - \lambda) Z_t \leq 0.
\end{aligned}$$

where the controls are $t_1 \in [0, T^{FB}]$, $\nu \in \Delta([0, t_1])$, and $z_t \geq 0$ for all t. Note that z_t is now a control chosen directly rather than a state governed by bonuses as in the no-inspection case. Condition $G_{t_1} \geq 0$ represents the global incentive constraint that fund-diversion at all times be less attractive than experimentation. The constraint $\ell \leq 0$ determines the local incentive constraint, determined analogously to the constraint in the non-inspection benchmark.

Let ψ^X denote the co-state associated with state $X \in \{O, B, Z, F, G\}$. The endpoint conditions (6.20) in (Arutyunov et al., 2018, p. 139) imply

$$\begin{aligned}
&\psi_{t_1}^O = \xi \in \{0,1\}, \quad && \psi_0^B \leq 0, \quad && \psi_0^Z \leq 0, \quad && \psi_{t_1}^F \leq 0, \quad && \psi_{t_1}^G \geq 0, \\
& && \psi_0^B B_0 = 0 && \psi_0^Z Z_0 = 0 && \psi_{t_1}^F (1 - F_{t_1}) = 0, && \psi_{t_1}^G G_{t_1} = 0.
\end{aligned}$$

The Hamiltonian is given by

$$\begin{aligned}
&H\left(O, B, Z, F, G\ ;\ b, z, \nu\ ;\ \psi^O, \psi^B, \psi^Z, \psi^F, \psi^G\ ;\ t\right) = \\
&\quad (1-F_t)\left[\psi_t^O \frac{1-\mu_0}{1-\mu_t}(\mu_t \lambda(1-b_t) - \varphi) - \psi_t^B \lambda b_t + \psi^G(\mu_0 e^{-\lambda t} b_t - \varphi)\right] + \psi_t^B \lambda B_t + \psi_t^Z \lambda Z_t.
\end{aligned}$$

The switching function, capturing the effect of measure ν, is given by

$$\begin{aligned}
&Q\left(O, B, Z, F, G\ ;\ b, z, \nu\ ;\ \psi^O, \psi^B, \psi^Z, \psi^F, \psi^G\ ;\ t\right) \\
&= \psi_t^O \frac{1-\mu_0}{1-\mu_t}\left(-\kappa + V^0(\mu_t, z_t)\right) + \psi_t^F - \psi_t^Z z_t + \psi^G \mu_0 (e^{-\lambda t} - e^{-\delta t}) z_t.
\end{aligned}$$

Henceforth, we will write $H(t)$ and $Q(t)$ and omit the other arguments unless a dependence shall be made explicit.

Let η_t be the multiplier associated with the mixed constrained ℓ. For each state X, the evolution of co-state ψ^X is governed by $d\psi_t^X = -\frac{\partial H}{\partial X}dt - \frac{\partial Q}{\partial X}d\nu_t + \frac{\partial \ell}{\partial X}\eta(t)dt$. As O_t and G_t, do not appear in H nor Q, their co-states are constant, i.e. $\psi_t^O = \xi \in \{0,1\}$ and $\psi_t^G \equiv \psi^G \geq 0$ for all t.

Any optimal path satisfies

$$\begin{aligned}&\frac{\partial H}{\partial b} + \frac{\partial Q}{\partial b} = \frac{\partial \ell}{\partial b}\eta_t\\ &-\xi(1-F_t)\mu_0 e^{-\lambda t}\lambda - (1-F_t)\psi_t^B\lambda + \psi^G\mu_0(1-F_t)e^{-\lambda t}\lambda = -\lambda(1-F_t)\eta_t,\end{aligned}$$

which simplifies to

$$\eta_t = \xi\mu_0 e^{-\lambda t} + \psi^G\mu_0 e^{-\lambda t} + \psi_t^B. \tag{3.A.2}$$

The co-states ψ_t^B and ψ_t^Z satisfy the differential relation

$$d\psi_t^B = -\lambda\psi_t^B dt + \lambda\eta_t dt \qquad \text{and} \qquad d\psi_t^Z = -\lambda\psi_t^Z dt + (\lambda-\delta)\eta_t dt.$$

Define the current value expressions for ψ_t^B as $\tilde{\psi}_t^B \equiv \frac{\psi_t^B}{e^{-\lambda t}}$ and likewise for ψ_t^Z and η_t. Then, we get

$$\begin{aligned}&\tilde{\psi}_t^B - \tilde{\eta}_t = \mu_0(\xi + \psi^G),\\ &d\tilde{\psi}_t^B = \tilde{\eta}_t\lambda dt,\\ &d\tilde{\psi}_t^Z = \tilde{\eta}_t(\lambda-\delta).\end{aligned}$$

Further, by (3.A.2) $d\tilde{\eta}_t = d\tilde{\psi}_t^B = \tilde{\eta}_t\lambda dt$, which leads to

$$\tilde{\eta}_t = \eta_0 e^{\lambda t}, \quad \tilde{\psi}_t^B = \psi_0^B - \eta_0 + \eta_0 e^{\lambda t} = \eta_0 e^{\lambda t} - \mu_0(\xi+\psi^G), \quad \tilde{\psi}_t^Z = \psi_0^Z + \eta_0(\lambda-\delta)\frac{e^{\lambda t}-1}{\lambda},$$

or in non-discounted expression, as both will be used below:

$$\eta_t = \eta_0, \qquad \psi_t^B = \eta_0 - \mu_0 e^{-\lambda t}(\xi + \psi^G), \qquad \psi_t^Z = \psi_0^Z e^{-\lambda t} + \eta_0(\lambda-\delta)\frac{1-e^{-\lambda t}}{\lambda}.$$

First, we show that $\xi = 1$, i.e. that the problem is not degenerate. Suppose to the contrary that $\xi = 0$. Then, for $\max_z Q(z)$ to be non-positive, we need $\psi_t^Z \geq \psi^G\mu_0(e^{-\lambda t} - e^{-\delta t})$ for all t and, in particular, $\psi_0^Z = 0$. This implies $\psi^G = 0$, which we show separately for the two cases $\delta > \lambda$ and $\delta < \lambda$.

In the case of $\delta > \lambda$, ψ_t^Z is decreasing, and we need for all t, that $\psi_t^Z = -(\delta-\lambda)\frac{1-e^{-\lambda t}}{\lambda} \geq \psi^G\mu_0(e^{-\lambda t} - e^{-\delta t})$. As both are 0 at $t = 0$, this requires that

$$\lim_{t\searrow 0}\frac{\partial}{\partial t}\left((\delta-\lambda)\frac{1-e^{-\lambda t}}{\lambda} + \psi^G\mu_0(e^{-\lambda t} - e^{-\delta t})\right) \leq 0.$$

This is equivalent to $\delta - \lambda + \psi^G\mu_0(\delta-\lambda) \leq 0$, which gives a contradiction as it requires $\delta \leq \lambda$.

In case $\delta \leq \lambda$, consider $\psi_t^Z = \eta_0(\lambda-\delta)\frac{1-e^{-\lambda t}}{\lambda} \geq \psi^G\mu_0(e^{-\lambda t} - e^{-\delta t})$. By (3.A.2) at $\psi_0^Z = 0$, we have $\eta_0 \leq \psi^G\mu_0$, inserting $\eta_0 \leq \psi^G\mu_0$ on the left-hand side of the inequality and considering the derivative of both sides with respect to t at t close to 0 show that this can only hold if $\psi^G = \eta_0 = 0$. Hence, $\psi^G = 0$ also if $\delta \leq \lambda$.

If $\psi^G = 0$, then $\eta_0 = \psi_0^B$, which implies that both are $= 0$ as $\eta_0 \geq 0$ and $\psi_0^B \leq 0$. This cannot be part of an optimal solution as all co-states are 0. Therefore, the problem does not degenerate, and we have $\xi = 1$.

Proof of Proposition 1

For exposition, wages were excluded from the optimal-control specification above. To see that they are 0 everywhere optimally, consider a contract with fixed inspection cdf F which is not everywhere equal to 0 and some incentive compatible combination of payment and continuation utility processes $(\tilde{b}, \tilde{w}, \tilde{z})$. The agent's expected utility is

$$\int_0^{t_1} \frac{1-\mu_0}{1-\mu_t} \left[(1-F_t)\left(\mu_t \lambda \tilde{b}_t dt + d\tilde{w}_t\right) + \mu_t \tilde{z}_t dF_t\right] \equiv \mu_0 u_0.$$

Determine the total expected pre-inspection wage payments $\tilde{P} = \int_0^{t_1}(1-F_t)\frac{1-\mu_0}{1-\mu_t}d\tilde{w}_t$ and consider the modified continuation value process z defined by

$$z_t = \tilde{z}_t + \frac{\tilde{P}}{\int_0^{t_1} \frac{1-\mu_0}{1-\mu_t}\mu_t dF_t}.$$

Substituting for u_0 in the objective, we get

$$-\mu_0 u_0 + \int_0^{t_1} \frac{1-\mu_0}{1-\mu_t}\left[(1-F_t)(\mu_t\lambda - \varphi)dt + \left(-\kappa + V^0(\mu_t, z_t) + \mu_t z_t\right) dF_t\right] dt.$$

We have $\frac{\partial}{\partial z}V^0(\mu, z) \geq -\mu$. Thus, for fixed value of u_0, increasing z_t increases the principal's payoff.

Finally the modification from $(\tilde{b}, \tilde{w}, \tilde{z})$ to $(\tilde{b}, 0, z)$ creates no profitable deviation. Consider the right hand side of the IC constraint:

$$\max_a \int_0^{t_1} (1-F_t)\left(\mu_0 e^{-\lambda\int_0^t a_s ds} + 1 - \mu_0\right)(1-a_t)\varphi dt + \mu_0 e^{-\lambda\int_0^t a_s ds} e^{-\delta\int_0^t (1-a_s)ds} z_t dF_t.$$

By construction of z, the expected utility without fund diversion remains equal:

$$\int_0^{t_1} \frac{1-\mu_0}{1-\mu_t}\left(-(1-F_t)d\tilde{w}_t + \mu_t(z_t - \tilde{z}_t)dF_t\right) = 0.$$

For general strategy a, the difference in expected utility resulting from the shift is

$$\int_0^{t_1}\left[-(1-\mu_0)(1-F_t)d\tilde{w}_t + \mu_0 e^{-\lambda t}\left(-(1-F_t)\lambda e^{\lambda\int_0^t(1-a_s)ds}d\tilde{w}_t + e^{-(\delta-\lambda)\int_0^t(1-a_s)ds}(z_t - \tilde{z}_t)dF_t\right)\right].$$

For all t, we have that $1 - \mu_0 + \mu_0 e^{\lambda\int_0^t(1-a_s)ds} \geq \mu_0 e^{(\lambda-\delta)\int_0^t(1-a_s)ds}$ for any action profile a. As the modification satisfies $d\tilde{w}_t \geq 0$ and $(z_t - \tilde{z}_t) \geq 0$, the integral must be smaller than the previous one. Hence, shifting wage payments to the future does not decrease the principal's payoff while decreasing all diversion incentives.

Now, consider the second part of the proposition, that is, $\lambda > \delta$. The (necessary) local condition for experimentation at time t was given by $\lambda b_t(1-F_t) \geq \frac{\varphi}{\mu_t}(1-F_t) + \lambda B_t + (\lambda-\delta)Z_t$. As B_t and Z_t are nonnegative, the necessity of positive bonus payments follows immediately in the case $\lambda > \delta$.

Finally, let $\delta > \lambda$. We have by (3.A.2) that $\eta_0 = \mu_0(1+\psi^G) + \psi_0^B$, which gives $\psi_t^Z = \psi_0^Z e^{-\lambda t} + (\mu_0(1+\psi^G) + \psi_0^B)(\lambda-\delta)\frac{1-e^{-\lambda t}}{\lambda}$.

Inserting into $Q(t)$ gives:

$$\begin{aligned}Q(t) &= \frac{1-\mu_0}{1-\mu_t}\left(-\kappa + V^0(\mu_t, z_t)\right)\\ &+ \left(\psi^G\mu_0(e^{-\lambda t} - e^{-\delta t} + (\delta-\lambda)\frac{1-e^{-\lambda t}}{\lambda}) - \psi_0^Z e^{-\lambda t} + (\mu_0 + \psi_0^B)(\delta-\lambda)\frac{1-e^{-\lambda t}}{\lambda})\right) z_t\\ &+ \psi_t^F.\end{aligned}$$

We can see that, if there is an inspection at some time $t > 0$, for Q(t) not to explode in z, the derivative of the bracket cannot be too high. In particular, if δ becomes arbitrarily large, $\psi_0^B < 0$ is required to maintain $\arg\max_z Q(t,z) \leq 0$. However, $\psi_0^B < 0$ implies $B_0 = 0$. No bonuses can be paid.

□

Proof of Proposition 2

The remaining contract duration after inspection at t, which we denote by $T(t) - t$, is determined by the first-order condition on Q with respect to z:

$$\frac{\partial Q}{\partial z} = \frac{1-\mu_0}{1-\mu_t}\frac{\partial V^0}{\partial z} + \left(\psi^G(e^{-\lambda t} - e^{-\delta t})\mu_0 - \psi_t^Z\right).$$

The benchmark deadline is determined by z such that $\frac{\partial V^0}{\partial z} = 0$. As V^0 is concave in z, the optimal value after inspection at t is larger than the benchmark if and only if the bracketed term above is positive. Substituting for ψ_t^Z gives:

$$\psi^G(e^{-\lambda t} - e^{-\delta t})\mu_0 - \psi_t^Z = \psi^G(e^{-\lambda t} - e^{-\delta t})\mu_0 - \psi_0^Z e^{-\lambda t} + \eta_0(\delta - \lambda)\frac{1 - e^{-\lambda t}}{\lambda}.$$

If $\delta - \lambda > 0$, the term is positive since $\psi_0^Z \leq 0$ and $\psi^G, \eta_0 \geq 0$. □

Proof of Proposition 3

To determine the inspection policy, we have to consider the shape of the switching function Q. Optimality requires that $Q(t) \leq 0$ for all t and $Q(t) = 0$ for all t in the support of the inspection distribution. The co-state ψ_t^F satisfies the differential relation

$$\begin{aligned} d\psi_t^F &= \xi\frac{1-\mu_0}{1-\mu_t}(\mu_t\lambda(1-b_t) - \varphi)dt - \psi_t^B\lambda b_t dt + \psi^G(\mu_0 e^{-\lambda t}b_t - \varphi)dt + \eta_t\left(\lambda b_t - \frac{\varphi}{\mu_t}\right)dt \\ &= \xi\frac{1-\mu_0}{1-\mu_t}(\mu_t\lambda - \varphi)dt - \psi^G\varphi dt - \eta_t\frac{\varphi}{\mu_t}dt. \end{aligned}$$

The second equality follows from (3.A.2) and conveniently implies that $Q(t)$ can be expressed independently of the bonuses. Integrating gives:

$$\psi_t^F = \psi_0^F + \mu_0\frac{1-e^{-\lambda t}}{\lambda}(\lambda - \varphi) - (1-\mu_0)\varphi t - \psi^G\varphi t - \eta_0\varphi t - \eta_0\frac{1-\mu_0}{\mu_0}\frac{e^{\lambda t} - 1}{\lambda}.$$

Consider $Q(t)$ after plugging in for ψ_t^F from above, z_t from (3.1), and V^0 from (3.A.1). After some algebra, this leads to:

$$\begin{aligned} Q(t) = &-\frac{1-\mu_0}{1-\mu_t}\kappa + (\psi^G\mu_0(e^{-\lambda t} - e^{-\delta t}) - \psi_t^Z - \mu_0 e^{-\lambda t})\varphi\left((T(t) - t) + \frac{1-\mu_0}{\mu_0}\frac{e^{\lambda T(t)} - e^{\lambda t}}{\lambda}\right) \\ &+ \mu_0\frac{1-e^{-\lambda T(t)}}{\lambda}(\lambda - \varphi) - (1-\mu_0)\varphi T(t) \\ &+ \psi_0^F - \psi^G\varphi t - \eta_0\varphi t - \eta_0\frac{1-\mu_0}{\mu_0}\frac{e^{\lambda t} - 1}{\lambda}. \end{aligned}$$

Differentiation w.r.t T gives the FOC for continuation utility after inspection at t:

$$0 = (\psi^G\mu_0(e^{-\lambda t} - e^{-\delta t}) - \psi_t^Z - \mu_0 e^{-\lambda t})\varphi\left(1 + \frac{1-\mu_0}{\mu_0}e^{\lambda T(t)}\right) + \mu_0 e^{-\lambda T(t)}(\lambda - \varphi) - (1-\mu_0)\varphi.$$

The above can be re-arranged as:

$$0 = (1-\mu_0)\frac{\mu_{T(t)}}{1-\mu_{T(t)}}\left(\mu_{T(t)}\lambda - \varphi\right) + \left(\psi^G \mu_0(e^{-\lambda t} - e^{-\delta t}) - \psi_t^Z - \mu_0 e^{-\lambda t}\right)\varphi.$$

Inserting ψ_t^Z:

$$(1-\mu_0)\frac{\mu_{T(t)}}{1-\mu_{T(t)}}\left(\mu_{T(t)}\lambda - \varphi\right) \tag{3.A.3}$$

$$+ e^{-\lambda t}\varphi\left(\psi^G \mu_0(1-e^{-(\delta-\lambda)t}) - \psi_0^Z + \eta_0(\delta-\lambda)\frac{e^{\lambda t}-1}{\lambda} - \mu_0\right) = 0. \tag{3.A.4}$$

The first bracket is positive if and only if $T(t) \leq T^{FB}$. Hence, the second bracket must be negative for the optimal duration $T(t)$ not to exceed T^{FB}. Furthermore, the second bracket is independent of $T(t)$ which implies that, if it were positive, the optimal $T(t)$ would be infinite and $Q > 0$, contradicting the optimality conditions. Therefore $T(t) \leq T^{FB}$ always.

If $\delta > \lambda$, the second bracket is smallest among all feasible co-state values when $\psi^G = \psi_0^Z = 0$ (Recall that we have $\psi^G \geq 0$ and $\psi^G \leq 0$). This gives an upper bound on t_1:

$$t_1 \leq \frac{1}{\lambda}\log\left(\frac{\mu_0\lambda}{\eta_0(\delta-\lambda)} + 1\right).$$

Further, the second bracket in (3.A.3) is smallest when $\eta_0 = 0$. This allows us to conclude that $\psi_0^Z \in [-\mu_0, 0]$. Similarly, we need $\psi^G \leq \mu_0 + \psi_0^Z$ for the bracket to be negative at $t = 0$.

Differentiating Q with respect to t gives:

$$\begin{aligned}
\frac{\partial Q}{\partial t} &= \mu_0 e^{-\lambda t}\lambda\kappa \\
&- \left(\psi^G \mu_0(e^{-\lambda t} - \frac{\delta}{\lambda}e^{-\delta t}) - e^{-\lambda t}((\psi_0^Z + \mu_0) + \eta_0\frac{\lambda-\delta}{\lambda})\right)\lambda\varphi\left((T(t)-t) + \frac{1-\mu_0}{\mu_0}\frac{e^{\lambda T(t)} - e^{\lambda t}}{\lambda}\right) \\
&- \left(\psi^G \mu_0(e^{-\lambda t} - e^{-\delta t}) - (\psi_0^Z e^{-\lambda t} + \mu_0 e^{-\lambda t} + \eta_0\frac{\lambda-\delta}{\lambda}(1-e^{-\lambda t})) + \eta_0\right)\varphi\left(1 + \frac{1-\mu_0}{\mu_0}e^{\lambda t}\right) \\
&- \psi^G\varphi,
\end{aligned}$$

which can be simplified to:

$$\begin{aligned}
\frac{\partial Q}{\partial t} &= \mu_0 e^{-\lambda t}\lambda\kappa \\
&- \left(\psi^G \mu_0(1 - \frac{\delta}{\lambda}e^{-(\delta-\lambda)t}) - \psi_0^Z - \mu_0 + \eta_0\frac{\delta-\lambda}{\lambda}\right)\lambda\varphi\left(e^{-\lambda t}(T(t)-t) + \frac{1-\mu_0}{\mu_0}\frac{e^{\lambda(T(t)-t)} - 1}{\lambda}\right) \\
&- \left(\psi^G \mu_0(1-e^{-(\delta-\lambda)t}) - \psi_0^Z - \mu_0 - \eta_0\frac{\delta-\lambda}{\lambda} + \psi^G e^{\lambda t}\mu_t\right)e^{-\lambda t}\frac{\varphi}{\mu_t}.
\end{aligned}$$

From the FOC on Q with respect to T (3.A.3), we know that

$$\psi^G \mu_0(1-e^{-(\delta-\lambda)t}) - \psi_0^Z - \mu_0 + \eta_0(\delta-\lambda)\frac{e^{\lambda t}-1}{\lambda} \leq 0,$$

which implies that, if δ is sufficiently large, the time derivative of Q is positive. Therefore, as inspection can happen only for times t with $Q(t) = 0$, we can conclude that there is a bound $\bar{\delta}$ such that for any $\delta \geq \bar{\delta}$, there can be at most one inspection time. That performing the inspection is optimal unless κ is too high, follows from comparing the optimal value without inspection given in the previous subsection with the value attainable with the inspection. At low κ, the latter is higher. □

References

Arutyunov, A., Karamzin, D. and Pereira, F. L. (2018). *Optimal Impulsive Control: The Extension Approach*, Vol. 477 of *Lecture Notes in Control and Information Sciences*, Springer.

Bergemann, D. and Hege, U. (1998). Venture Capital Financing, Moral Hazard, and Learning, *Journal of Banking & Finance* **22**(6-8): 703–735.

Bergemann, D. and Välimäki, J. (2008). *Bandit Problems, The New Palgrave Dictionary of Economics*, 2nd edn, Macmillan Press.

DeJarnette, P., Dillenberger, D., Gottlieb, D. and Ortoleva, P. (2020). Time Lotteries and Stochastic Impatience, *Econometrica* **88**(2): 619–656.

Gompers, P. A. and Lerner, J. (2004). *The Venture Capital Cycle*, MIT press.

Green, B. and Taylor, C. R. (2016). Breakthroughs, Deadlines, and self-reported Progress: Contracting for multistage projects, *American Economic Review* **106**(12): 3660–99.

Halac, M., Kartik, N. and Liu, Q. (2016). Optimal Contracts for Experimentation, *The Review of Economic Studies* **83**(3): 1040–1091.

Hörner, J. and Samuelson, L. (2013). Incentives for Experimenting Agents, *RAND Journal of Economics* **44**(4): 632–663.

Keller, G., Rady, S. and Cripps, M. (2005). Strategic Experimentation with Exponential Bandits, *Econometrica* **73**(1): 39–68.

4
Dynamic Incentives with Costly Inspections

Joint with Peter A. Wagner

4.1 Introduction

Critics and advocates alike recognise the European Union's potential as a 'regulatory superpower' (The Economist, 2020). Few manufacturers can afford not to comply with EU regulation and thereby renounce its single market with more than 500m consumers. The EU prides itself on passing demanding standards so that abiding by its regulation is sufficient to fulfil most non-EU countries' requirements. A 'Rise of the Regulatory State in Europe' was already observed by Majone (1994), referring to the shift in state interventions away from public ownership and centralised planning to more fine-tuned regulation.

Regulation with fine-tuned interventions requires effective procedures to ensure compliance with standards for production and accounting procedures or transactions which are often complex and hard to observe. Recognising the challenges of compliance, the European Commission provides substantial guidance in the form of non-binding recommendations, best-practice benchmarks, and case studies. A case in point is the recommendation (EU) 2019/138 (European Commission, 2019) which 'provides a framework to help exporters identify, manage and mitigate risks associated with dual-use trade controls and to ensure compliance with the relevant EU and national laws and regulation.' Dual-use goods are products with civil as well as military applications that fall under special regulation to promote international security, e.g. by 'countering risks associated with the proliferation of Weapons of Mass Destruction' (European Commission, 2019, p. 17). The recommendation focuses on the established practice of elaborating an *Internal Compliance Programme* (ICP). The essential components of ICPs can be summarised with the following quotes.

> A well-functioning ICP has clear **reporting** procedures about the notification and escalation actions of employees when a suspected or known incident of non-compliance has occurred. As part of a sound compliance culture, employees must feel confident and reassured when they raise questions or report concerns about compliance in good faith. Performance reviews, **audits** and reporting procedures are designed to detect inconsistencies to clarify and revise routines if they (risk to) result in non-compliance.
>
> identifies and appoints the person(s) with the overall responsibility to ensure the corporate compliance commitments.The internal organisational structure (European Commission, 2019, pp. 20 & 25)

Further, the document suggests that the person(s) responsible for compliance 'should have the **power** to stop transactions', and to establish a 'set of **remedial actions** to guarantee the proper implementation of the ICP' (European Commission, 2019, p. 24).

The aim of this paper is to study how these elements, *reporting* procedures, *audits*, and potential *punishments* are optimally combined to ensure maximal compliance in various situations. Is it possible to implement full compliance and, if so, what are the incentives of a compliance manager to perform inspections of which he already knows the outcome?

To capture the most salient ingredients of this problem, we present a dynamic game between a principal and an agent. The principal is tasked with ensuring that the agent complies with regulation. The agent privately observes the current state of compliance which is subject to changes at random times. Whenever a potential change arrives, the agent's current effort determines the probability of maintaining or obtaining compliance. At each point in time, the agent chooses privately how much effort to invest. He then observes the compliance state and makes a report to the principal. The principal decides whether to verify this report at a cost and may choose to punish the agent. Further applications captured by our model include the following:

Risk-management in banks. Applying our model to the financial sector, the principal represents the authority aiming to ensure that banks apply proper risk assessment.[1] The agent represents a financial institution. The banking authority can verify the bank's risk exposure by conducting stress-tests or on-site audits, and it has the power to punish banks either through monetary fines, by demanding changes

[1] In EU member states, this is the responsibility of the central bank and/or the finance ministry.

in the management structure, or by denying the bank license. We revisit this application in Section 4.5 where we discuss possible means to enable randomisation without commitment.

General Data Protection Rules (GDPR). Since May 2018, the processing and handling of private data of users in the European Union falls under a EU-wide set of rules (EU, 2016). With particular emphasis on the online context, these rules are especially relevant for technology firms as those pooled under the acronym GAFAM.[2] The regulation requires data-handling organisations to identify a responsible *data protection officer*. In our model, the principal represents the data protection officer who is tasked with ensuring and monitoring compliance by the organisation's staff with the regulation.

Results. Our main result shows how full compliance can be attained in equilibrium requiring minimal inspection costs from the principal. This equilibrium entails two phases that depend on the agent's report: a penalty phase and a monitoring phase. Reporting non-compliance induces the penalty phase in which the agent pays a fixed flow fine, but is not inspected. Reporting compliance induces the monitoring phase in which the agent is never fined but subject to periodic inspections. The dynamic incentives in our setting require a transition fine, that is, a penalty the agent faces for reporting a failure in compliance while in the monitoring phase. The fine increases as the next inspection approaches. The transition fine is needed because the agent would otherwise have an incentive to delay reporting instances of non-compliance in the monitoring phase, in the hope to recover compliance prior to the next inspection.

To characterise this optimal equilibrium, we find a lower bound on the inspection costs required to provide incentives for compliance and then show how this bound can be attained. This bound is established by two results. First, the principal can achieve compliance at the same costs as if she were able to commit to any *predictable* inspection protocol. Second, a non-committed principal cannot benefit from randomisation. Thus, we can identify an optimal equilibrium by solving a related mechanism design problem in which the principal can commit to a cost-minimising predictable inspection protocol subject to the agent's incentive compatibility constraints.

Comparative statics reveal that compliance can be achieved at lower inspections costs when the maximal punishment the principal can impose is larger or when the agent's effort cost is smaller. The arrival rate of transitions and the connection between effort and realised compliance *conditional* on arrival have ambiguous effects on the inspection costs. In both cases, an increase gives the agent more control over

[2] Google, Apple, Facebook, Amazon, Microsoft.

the state. This makes the current state of compliance more informative about effort shortly before but erodes the link between current compliance and effort further in the past. We show that a high degree of control for the agent makes incentive provision through inspections arbitrarily costly. This cost increase stems from the fact that, without commitment, inspections cannot be randomised, so the agent is confident to improve quality prior to inspection whenever it lies in the strict future.

The cost increase is closely related to the value of randomised inspections. Considering the contracting problem for a committed principal without the restriction to predictable inspections we show that a random mechanism dominates predictable plans. In light of our result that a non-committed principal cannot benefit from randomisation, we discuss possible sources of commitment power or incentives to inspect. In particular, we conclude that the separation of inspection planning and execution as observed in banking supervision is a promising measure.[3]

Our results shed light on the role of commitment in effective compliance procedures. Commitment power is important for costly inspections in two fundamental and intertwined ways: First, when full compliance is expected from the agent, the principal has no incentive to pay inspection costs to reveal the agent's private information. For instance, Reinganum and Wilde (1985) show that full compliance is not achievable without commitment in static games. Second, randomised inspections – which may be more effective in providing incentives – are more demanding on the principal's commitment power.

With repeated interactions in which the principal's choices affect the continuation play, there is scope to provide punishment for insufficient inspection. Indeed, Ben-Porath and Kahneman (2003) prove a folk-theorem, showing that full compliance can be obtained without commitment in the undiscounted limit through public reports and random inspections. In our game, full compliance is attainable even with discounting. However, the non-committed principal cannot lower inspection costs through randomisation. These differences stem from the observability of inspections. In Ben-Porath and Kahneman (2003), inspections are not observable so that some instances of non-compliance are required to identify and incentivise inspections. When these incentives are provided tightly, that is, such that the principal is indifferent between inspecting or not, and inspections are unobserved, the principal may as well randomise. In the equilibrium constructed in Ben-Porath and Kahneman (2003), the frequency of non-compliance vanishes as the discount factor approaches one. In the applications we want to study with this paper, it seems sensible to assume that the agent observes inspections carried out by the principal. This makes it possible to provide strict incentives for the principal as the failure to inspect can be punished. Randomisation, however, requires the principal to be

[3] See discussion in Section 4.5.

indifferent between inspecting or not. We show that this precludes benefiting from randomisation without commitment power. This may explain the findings in Chang et al. (1993), demonstrating a high degree of predictability of Medicare audits in the US.

Related literature. The literature on costly state verification (CSV) has been influential in explaining debt contracts and the role of financial intermediaries. A number of papers have studied dynamic extensions of the static models by Townsend (1979), Gale and Hellwig (1985), Mookherjee and Png (1989) and Border and Sobel (1987). These include models with a risk-neutral agent and deterministic monitoring schemes (Webb, 1992, Chang, 1990), with a risk-neutral agent and random verification under wealth constraints (Monnet and Quintin, 2005, Antinolfi and Carli, 2015), and models with randomised verification and a risk-averse agent (Wang, 2005, Popov, 2016). These papers focus on discrete time settings in which the agent's information is i.i.d. across periods. More similar to the setup in this paper is Ravikumar and Zhang (2012) which also considers a continuous-time model with persistent private information. However, they consider a pure adverse selection framework in which private information results from a single exogenous jump in income that is unobservable to the principal. In contrast, we consider an information structure that allows for oscillations between states, and effort-dependent transition rates.

All of the above papers focus on optimal mechanisms in which the principal has full commitment power. Less attention has been given to the question of limited commitment in costly state verification models. A notable exception is the paper by Krasa and Villamil (2000) which considers an extension of the standard CSV model in which the principal is free to choose whether to go to court to enforce a contract. Their results are similar in spirit to ours, in that randomisation is optimal when the principal can commit ex-ante, but deterministic verification is optimal without commitment. However, they use a static model and rely on the assumption of an external party that enforces the contract. The strategic concerns in our dynamic model without third-party enforcement are somewhat different.

There is another strand of related literature that is concerned with deterrence of crimes and illegal behaviour through policing and punishment (Becker, 1968, Bassetto and Phelan, 2008, Bond and Hagerty, 2010, Dye, 1986). The primary focus of these papers is the enforcement of an agent's hidden action. This stands in contrast to the CSV literature, which focuses on the problem of eliciting hidden information. For the case of limited commitment, there is an extensive literature on so-called 'inspections games' which have been applied to various problems, among them pollution and arms control. The various contributions to this literature, including static as well as dynamic models, are surveyed by Avenhaus et al. (2002). A feature of equilibria in inspection games is that without commitment, it is impossible to induce the agent to comply fully, which stands in contrast with the results in this paper.

The basic modelling of state transitions is based on Board and Meyer-ter-Vehn (2013) which studies the reputation problem of a firm that has to make costly investments towards quality improvements when quality becomes publicly observable at random times. We embed this framework into a principal-agent problem, allowing the principal to control the agent's payoff and to choose when to verify quality. A similar model has been studied by Kim (2015) in the context of environmental control. This author focuses on specific classes of inspection policies without limited liability and does not solve for the optimal contract. Most closely related is the independent contribution by Varas et al. (2020). These authors also study optimal monitoring policies in a principal-agent model with the same inspection technology, but the incentive problem is somewhat different. In their model, inspections make the agent's type public. They focus on mechanisms without reporting and without transfers where the agent is motivated by the desire to generate a positive reputation in the market place at future inspection dates.

Our model is also related to the *machine maintenance problem* in operations research. The machine maintenance problem is a statistical decision problem in which a machine 'fails' at random times which can be observed only through inspection (see Osaki, 2002, for an overview). Similar models have also been applied in the accounting literature to study the optimal timing of audits (Kaplan, 1969, Carey and Guest, 2000, Hughes, 1977).

4.2 Model

4.2.1 Preliminaries

There are an agent and a principal. Time $t \in [0, \infty)$ is continuous. The principal and the agent are risk-neutral and discount future payoffs at a common rate $r > 0$. The agent is required to comply with exogenously given regulation. Compliance is represented by a binary indicator variable $\theta_t \in \{0, 1\}$ which fluctuates over time. At each instant, the agent chooses effort level $\eta_t \in [0, 1]$ at instantaneous cost of $c\eta_t dt$. Effort affects transition rates of compliance as follows.

State dynamics. Let $(\Omega, \mathcal{F}, P)$ be a probability space, and let $\{\mathcal{F}_t\}$ be a filtration of the sigma-algebra $\mathcal{F}$. Let the marked point process $z = \{z_t\}_{t\geq 0}$ represent the arrival of random shocks, where $z_t = 0$ except at isolated times $t_0 < t_1 < \ldots$ arriving at constant rate λ. At each random time t_j, the value of the shock z_{t_j} is uniformly distributed on $[0, 1]$. Let $\{\mathcal{F}_t\}$ be the natural filtration generated by z. The process $\theta = \{\theta_t\}_{t\geq 0}$ evolves according to a two-state Markov process that depends on the shock process and the agent's effort choices. Immediately after the arrival of a shock z_t, $\theta_t = 1$ if $\alpha\eta_t \geq z_t$, and $\theta_t = 0$ if $\alpha\eta_t < z_t$. The factor $\alpha \in (0, 1)$ is a noise factor that represents the possibility that the agent cannot always maintain compliance despite best efforts. At any time t, the probability of $\theta_{t+dt} = 1$ is therefore

$$\text{Prob}(\theta_{t+dt} = 1|\theta_t) = \begin{cases} 1 - \lambda dt + \alpha\eta_t\lambda dt & \text{if } \theta_t = 1, \\ \alpha\eta_t\lambda dt & \text{if } \theta_t = 0. \end{cases}$$

Monitoring and fines. The principal has an interest in the agent's compliance. At each instant t, the principal's flow payoff is $\theta_t H$ where $H > 0$. The agent privately observes θ_t at each time $t \geq 0$. The principal cannot observe the agent's effort and compliance is observable only through costly inspections. At any in time $t \geq 0$, the principal can investigate the agent at lump-sum cost $\kappa > 0$. In addition to inspection decisions, the principal can punish the agent through fines. Fines may be understood literally, as compulsory monetary payments, or they may be interpreted as remedial actions, for example as mentioned in European Commission (2019), that negatively impact the agent in some other way. This may include vetoing export transactions or denying banking licenses. We assume that the principal does not benefit from fining the agent directly. This assumption prevents rent-seeking incentives for the principal, who could use fines as a means to transfer surplus. In the context of public institutions, this assumption represents a benevolent view of government that uses transfers with the intention to correct market failures.

Both players are allowed to 'exit', which permanently ends the relationship and results in a continuation value of zero for the principal, and continuation payoff of $-B$ for the agent. For the principal, this implies a constraint on the severity of the fines she can impose. We assume that the exogenously given bound B is larger than $\frac{c(r+\lambda)}{\alpha\lambda r}$. Otherwise, the maximal punishment is insufficient to incentivise effort. The players' option to exit reflects the idea that they can limit their liability by dissolving the relationship: a compliance manager or employee can quit her his, a firm or bank can shut down, etc.

Timing. The timing at each $t \geq 0$ is as follows.[4] First, the agent chooses effort level η_t. Subsequently, nature determines whether a technology shock arrives and, conditional on the arrival of a shock and the effort level η_t, draws a new compliance level. The agent then observes the realised state and sends a report $\hat{\theta}_t \in \{0, 1\}$ to the principal. The principal, in turn, makes an inspection decision and, conditional on the inspection outcome, chooses a fine immediately incurred by the agent. Denote

[4]We outline the sequentiality at a given instant to establish some intuition about the order of events over time. Formally, this order is captured by continuity properties of the respective action and state paths.

by N_t^I the number of inspections and by F_t the cumulative fines up to and including time t.

Histories and strategies. A history at time t is a collection of paths

$$h_t = \{\eta_s, \theta_s, \hat{\theta}_s, N_s^I, F_s\}_{s\in[0,t]},$$

where

$$(\eta_s, \theta_s, \hat{\theta}_s, N_s^I, F_s) \in [0,1] \times \{0,1\} \times \{0,1\} \times \mathbb{N}_0 \times \mathbb{R}_+.$$

Throughout, we denote by subscript $t-$ strict histories h_{t-} for which the realisation at time t are excluded. Let H_t be the set of all time-t histories and H_{t-} the set of all strict histories. Let $H = \bigcup_{t\geq 0} H_t$ and $H_- = \bigcup_{t\geq 0} H_{t-}$.

The agent's strategy specifies efforts and reports as functions of histories. A strategy for the agent is then defined as a pair $(e,r) = (\{e_t, r_t\}_{t\geq 0})$ with

$$e_t : H_{t-} \to [0,1], \quad r_t : H_{t-} \times \{0,1\} \to \{0,1\},$$

where $e_t(h_{t-})$ is the agent's effort at time t and $r_t(h_{t-}, \theta_t)$ is the agent's report at time t after history h_{t-} when the state at time t is θ_t.

To capture the principal's uncertainty about the agent's effort choices and the true level of compliance, we consider a partition $\mathcal{H}_t^P$ of the history set H_t at any t which comprises all subsets of H_t that are indistinguishable to the principal. Define the partition $\mathcal{H}_{t-}^P$ similarly for strict histories at t. To allow for randomised inspections, we assume the principal is equipped with a (private) random signal π, defined on a sufficiently rich probability space with state space Π. A strategy for the principal is defined as a pair $(n,f) = (\{n_t, f_t\}_{t\geq 0})$ of mappings

$$n_t : \Pi \times H_{t-} \times \{0,1\} \to \{0,1\}, \quad f_t : H_{t-} \times \{0,1\}^3 \to \mathbb{R}_+,$$

which are constant on every $H_{t-}^P \in \mathcal{H}_{t-}^P$ for each $t \geq 0$. Here, $n_t(\psi, h_{t-}, \theta_t)$ is equal to 1 if an inspection is performed at time t and equal to 0 otherwise, and by $f_t(h_{t-}, \theta_t, \hat{\theta}_t, dN_t^I)$ we denote the fine imposed by the principal at time t. We abuse notation slightly and write $f_t(h_t)$ instead of $f_s(h_{t-}, \theta_t, \hat{\theta}_t, dN_t^I)$ whenever there is no danger of confusion.

The exit decision for each player at any history is a binary variable indicating whether this player decides to exit or not. For the ease of exposition, we do not introduce additional notation for these choices. The strategies above are to be understood as conditional on no player having exited previously. Actions to be chosen after one player exited are irrelevant.

Equilibrium. In continuous-time games with observable actions, strategies may not produce well-defined action paths and – in stochastic environments – agents' behaviour may be non-measurable. We adopt the approach by Kamada and Rao (2018) and impose restrictions on strategies to ensure well-defined action paths. We refer the interested reader to the appendix. We do not impose restrictions on strategies that rule out non-measurable behaviour. Instead, our equilibrium definition requires that strategies lead to measurable actions along the equilibrium path. Histories away from the equilibrium path may lead to non-measurability. Payoffs at such histories can be assigned freely within feasible bounds. In our game, these bounds can be reached unilaterally by either player at any history, so non-measurablilty off path cannot be used as a threat to enlarge the equilibrium set (see also the discussion of this approach in Kamada and Rao, 2018).

For a given path realisation $h = \{\eta_t, \theta_t, \hat{\theta}_t, N_t^I, F_t\}_{t\in[0,\infty)}$ the discounted net present payoff for the principal at time t is

$$v_t = \int_t^\infty e^{-r(s-t)} \left(\theta_s H ds - \kappa dN_s^I\right). \tag{4.1}$$

Similarly, the discounted net present payoff for the agent at time t is given by

$$u_t = \int_t^\infty e^{-r(s-t)} \left(-c\eta_s ds - dF_s\right). \tag{4.2}$$

Given a strategy profile, the principal and the agent form expectations about h based on their past observations whenever possible. For strategies that induce measurable action processes along the equilibrium path, we denote the expected payoff for the agent and the principal at t by $U_t(h_{t-}) = E[u_t|h_{t-}]$ and $V_t(h_{t-}) = E[v_t|h_{t-}]$ respectively, where the expectation is with respect to $\{z_s\}_{s\in(t,\infty)}$ and π.

We define a combination of strategies $((e, r), (n, f))$, together with processes of expectation operators $\{V_t, U_t\}_{t\geq 0}$, to be a *perfect Bayesian equilibrium* if the following holds.

1. There is no alternative strategy for the principal that yields a strictly higher payoff at any t given her expectation.

2. There is no alternative strategy for the agent that yields a strictly higher payoff at any t given his expectation.

3. Along the equilibrium path, the payoffs V_t and U_t are equal to the conditional expectations given above. Away from the equilibrium path, V_t and U_t are equal to the conditional expectations whenever these are well-defined.

4. For all $h_{t-} \in H_{t-}$, we have $V_t(h_{t-}) \geq 0$, and $U_t(h_{t-}) \geq -B$.

4.3 Construction of the Principal-Optimal Equilibrium

The main objective in this paper is to construct equilibria that achieve maximum compliance of the agent at the lowest possible costs for the principal. We say that an equilibrium achieves maximum compliance if, after any history along the equilibrium path, the agent exerts maximum effort. In this way, the probability of attaining or remaining in compliance at any given time is maximised. We assume at this point that H is large, relative to the inspection cost κ, so that the value for the principal generated from maximum compliance outweighs the monitoring costs that is needed to compel the agent to exert effort. The exact bound on H will follow from the characterisation of the principal-optimal equilibrium which we present in Section 4.4.

Our construction of the principal-optimal equilibrium proceeds in two steps. We first show that sequential rationality for the principal is equivalent to the requirement that the inspection schedule be predictable for the agent.[5] The equilibrium optimisation is therefore equivalent to finding the cost-minimising predictable strategy for the principal. This is a mechanism design problem, which we then set up in recursive form and solve using dynamic programming techniques for piecewise deterministic processes (Davis, 1993). We show equivalence in the following subsection. The mechanism design problem is set up and solved in Subsection 4.3.2. Readers who prefer to skip the technical details can proceed directly to the equilibrium description in Section 4.4.

4.3.1 Sequential Rationality and Predictability of Inspections

It is sequentially rational for the principal to carry out inspections only if failing to do so results in some form of punishment. To provide such punishments for the principal, inspections must be at least partially predictable by the agent. In fact, the following result shows that the principal cannot gain from any non-predictability in the timing of inspections.

Lemma 1. *For any maximum compliance equilibrium, there exists a maximum compliance equilibrium with predictable inspections that generates the same expected payoff for the principal.*

The formal proofs of this and the remaining results are relegated to the appendix. The fact that the principal cannot gain from non-predictable inspections in equilibrium is due to the lack of commitment power. Indeed, the principal can do strictly

[5] Here, predictability means that inspections are measurable with respect to the information available to the agent, so that he knows at any history whether or not an inspection will take place. Henceforth, we refer to inspections as random whenever they are non-predictable for the agent.

better when she can commit to inspecting at random. We will revisit this issue and potential ways to mitigate the commitment problem in Section 4.5.

The idea behind the proof is to consider the worst realisation of the principal's mixed equilibrium strategy and argue that the outcome generated by this strategy can be replicated by a predictable strategy. More specifically, consider any maximum compliance equilibrium in which the principal randomises over inspection dates. For a random strategy to be optimal for the principal, she must be indifferent between all inspection processes consistent with her strategy, and at any later date among those processes that are consistent with past play. Call an inspection process *most vigilant* if, after any history, it generates the earliest possible inspection dates that are consistent with the underlying random strategy. From the agent's perspective, there is zero probability that the principal performs inspections any earlier than given by the most vigilant inspection process. Moreover, by construction, the agent's incentive-compatibility conditions must be satisfied between inspections. Therefore, we can replace the principal's random strategy with a strategy in which inspections are predictable and determined by the most vigilant inspection process. With this new strategy, the incentive-compatibility conditions for the agent continue to be satisfied. By indifference, the principal receives the same payoff as in the original equilibrium.

The next result shows that the predictability of inspections is indeed the only restriction implied by sequential rationality. That is, for any strategy combination with well-defined action paths, predictable inspections and maximum effort as the agent's best response, there exists a perfect Bayesian equilibrium that induces the same outcome.

Lemma 2. *Let $((n,f),(e,r))$ be a strategy profile and suppose the following holds.*

(i) The principal's strategy (n,f) is predictable.

(ii) The agent's strategy (e,f) is a best response to the principal's strategy (n,f), and $\eta_t = 1$ at all $t \geq 0$ for every action process generated by the profile $((n,f),(e,r))$.

(iii) The action path generated by the strategy profile $((n,f),(e,r))$ is measurable.

(iv) The expected payoff for the principal is non-negative along any history generated by $((n,f),(e,r))$.

Then there exists a perfect Bayesian equilibrium $((n^*,f^*),(e^*,r^*))$ which generates the same distribution over action paths as $((n,f),(e,r))$.

Intuitively, predictability makes it easy to incentivise the principal because the agent immediately detects when an inspection does not take place as anticipated.

The last two items in the lemma are needed to ensure that the strategy combination satisfies Conditions (3.) and (4.) in our equilibrium definition. By hypothesis, the value from compliance H is large enough so that for a principal-optimal strategy combination, item (iv) is automatically satisfied.

4.3.2 Derivation of the Principal-Optimal Predictable Strategy

The two lemmas of the previous subsection jointly imply that we can transform our equilibrium optimisation problem into a mechanism design problem in which inspections must be predictable for the agent. By the revelation principle, it is without loss to focus on direct mechanisms with truthful reporting as any indirect mechanism can be replicated with a direct mechanism by giving the agent the same payoff that he would have obtained by misreporting his information in the original mechanism. A direct mechanism is fully characterised by a strategy for the principal. To derive the optimal direct mechanism, we use the martingale representation theorem to map a given strategy of the principal into a law of motion for the *promised utilities* of the agent and represent the incentive-compatibility conditions as constraints on the evolution of promised utilities. Standard results then allow us to formulate the optimisation problem of the principal in recursive form, using promised utilities as state variables.

Promised Utilities and Incentive Compatibility

Begin by fixing an arbitrary strategy (n, f) for the principal and let U_t be the agent's expected discounted continuation payoff from t onward, assuming he exerts maximum effort and reports compliance truthfully throughout. Here, we consider the general case in which the principal is allowed to use random inspections, as we will need to refer back to these results later in Section 4.5. To proceed, define W_t to be the agent's lifetime expected utility, with expectations taken with respect to the information that is available up to time t:

$$W_t = \int_0^t e^{-rs} \left(-dF_s - c\eta_s ds\right) + e^{-rt} U_t$$

By construction, the process $\{W_t\}_{t\geq 0}$ is a martingale. Note that there are three types of random effects here: changes in compliance, changes in reports, and inspections. Inspections are governed by the process N^I given by the principal's strategy. For the sake of consistency, we also introduce the counting processes $N^\theta = \{N_t^\theta\}_{t\geq 0}$ and $N^{\hat{\theta}} = \{N_t^{\hat{\theta}}\}_{t\geq 0}$ that count the number of changes in the state of compliance and reports, respectively. For each process N^a with $a \in \{\theta, \hat{\theta}, I\}$, define the *compensator* to be a predictable process $\nu^a = \{\nu_t^a\}_{t\geq 0}$ such that the compensated process $N_t^a - \nu_t^a$ is a martingale. The compensator exists under very general conditions and can be interpreted as the predictable drift of the underlying (non-predictable) stochastic process. Alternatively, we can think of the compensator as a generalisation of the cumulative hazard function, and consequently of $d\nu^a/dt$ as the hazard rate of transitions in N_t^a (if it exists). The martingale representation theorem for marked point processes (Last and Brandt, 1995) implies the following result.

Lemma 3. *There exist predictable processes Δ^θ, $\Delta^{\hat{\theta}}$, Δ^I such that the evolution of the agent's promised utility is given by*

$$dU_t = rU_t dt + dF_t + c\eta_t dt + \sum_{a\in\{\theta,\hat{\theta},I\}} \Delta_t^a (dN_t^a - d\nu_t^a). \tag{4.3}$$

The processes Δ^θ, $\Delta^{\hat{\theta}}$ and Δ^I have an intuitive interpretation: Δ_t^θ represents the jump in utility that results from a change in compliance at time t. Similarly, $\Delta_t^{\hat{\theta}}$ is the jump in utility that results from a change in reported compliance and Δ_t^I represents the jump in utility that results from an inspection at time t.

The characterisation of the evolution of payoffs in Equation (4.3) enables us to write the mechanism design problem in recursive form, and therefore, we can use a dynamic programming approach to solve it. Because of the persistence in compliance, the state variable in our problem must keep track of two utility promises, one for each possible state (see Fernandes and Phelan, 2000). Intuitively, an additional state variable is needed to keep track of private valuations of future payoffs, since these differ depending on the current state and are not common knowledge at the time the agent chooses his report.

Formally, consider a principal's strategy which is incentive compatible with truthful reporting and maximum effort. Take a strict history of events at any time t (i.e. the history up to, but not including t). Now, define

$$\begin{aligned} U_t^0 &= \mathbb{E}_{t-}[U_t|\theta_t = 0], \\ U_t^1 &= \mathbb{E}_{t-}[U_t|\theta_t = 1] \end{aligned} \tag{4.4}$$

to be the promised utilities at h_t for each possible realisation of θ_t. Here $\mathbb{E}_{t-}$ represents the expectation conditional on all available information up to time t. Following Zhang (2009), we call U_t^1 the *persistent payoff* in state $\theta_{t-} = 1$, and the *transitional payoff* in state $\theta_{t-} = 0$, and vice versa for U_t^0. The state for the principal's contracting problem consists of three objects: the most recent state θ_{t-}, and the promised utilities U_t^0 and U_t^1. The following lemma provides a complete characterisation of the agent's incentive-compatibility constraints in terms of these variables.

In addition to delivering the utility the agent is promised (promise-keeping constraint), the contract has to provide incentives for the agent to truthfully reveal the state (honesty constraint), incentives for exerting maximum effort (obedience constraint) and it must be sufficiently lenient to deter the agent from withdrawing (participation constraint). By standard arguments, it is optimal for the principal to enforce the most severe punishment after a verified misreport to induce the agent to reveal the state truthfully. Thus, the agent's payoff after a false report that was detected by an inspection is $-B$.

Lemma 4. *A principal's strategy that generates the process $\{U_t^1, U_t^0\}_{t\geq 0}$ of promised utilities induces maximum effort and truthful reporting if and only if there exists a predictable process $\{d\mu_t \geq 0\}_{t\geq 0}$ such that for $i = \theta_t$ and $j = 1 - \theta_t$ we have*

(Pk) $dU_t^i = rU_t^i dt + \lambda(i-\alpha)(U_t^1 - U_t^0)dt + dF_t + cdt - d\nu_t^I \Delta_t^I.$

(H) $dU_t^j = rU_t^j dt + \lambda(j-\alpha)(U_t^1 - U_t^0)dt + d\nu_t^I(B + U_t^j) + dF_t + cdt - d\mu_t$

(O) $U_t^1 - U_t^0 \geq c/\lambda\alpha,$

(P) $U_t^0, U_t^1 \in [-B, 0]$

at all $t \geq 0$ with $dN_t^a = 0$ for each $a = \theta, \hat{\theta}, I$.

Condition (Pk) is the promise-keeping constraint which is the expectation of Equation (4.3) in Lemma 3 conditional on no intervention at time t. Condition (O) is the obedience constraint that ensures that the agent exerts maximum effort. The condition has a reasonably straightforward interpretation. The right-hand side of this inequality is the marginal cost of effort. The left-hand side represents the marginal gain from effort. The factor $\lambda\alpha\eta_t$ is the rate at which a shock generates compliance at effort η_t. The utility gain from compliance is $U_t^1 - U_t^0$. Therefore, $\lambda\alpha(U_t^1 - U_t^0)$ is the marginal gain from effort for the agent. Thus, (O) states that for maximum effort to be optimal for the agent, the marginal gain must exceed the marginal cost.

Condition (H) is the honesty constraint that ensures that the agent reports truthfully. This constraint says that transitional utility cannot increase too quickly. The variable $d\mu_t$ is the choice of the principal who can use it to lower transitional utility. We shall refer to it as the principal's *threat.* The variable $d\nu_t^I > 0$ can be interpreted as the rate of inspections. When an inspection reveals a misreport, the agent's continuation value for the agent is $-B$. Had he reported truthfully instead, he would have received transitional utility U_t^j. Thus, $-B - U_t^j$ is the promised utility the agent loses if an inspection reveals a false report. The honesty constraint (H) can be illustrated using heuristic arguments. For simplicity, we focus here on the specific case with $\theta_t = 0$ and without random inspections. Suppose a failure of compliance occurs at time $t \geq 0$. The agent is willing to report the decline without delay only if he can't gain from delaying the report. In particular, this means that on a small interval $[t, t + dt)$, the value of admitting non-compliance must exceed the value of misreporting compliance, i.e.

$$U_t^0 \geq \int_t^{t+dt} e^{-(r+\alpha\lambda)(s-t)}(\alpha\lambda U_s^1 ds - dF_s - cds) + e^{-(r+\alpha\lambda)dt} U_{t+dt}^0.$$

The integral is the instantaneous gain from reporting the high instead of the low state, followed by a truthful report of the true state at time $t + dt$ if no change happened in the meantime. Taking a first-order approximation, we obtain, after a few rearrangements,

$$U_t^0 \geq \alpha\lambda U_t^1 dt - dF_t - cdt + (1 - rdt - \alpha\lambda dt)U_{t+dt}^0.$$

If we further substitute the approximation $dU_t^0 := U_{t+dt}^0 - U_t^0$ and ignore higher order terms, then this necessary condition for a truthful report is equivalent to

$$dU_t^0 \leq rU_t^0 dt - \alpha\lambda(U_t^1 - U_t^0)dt + dF_t + cdt.$$

This inequality is precisely condition (H) in Lemma 4 for the high state without random inspections. Note that while this heuristic derivation generates a necessary condition, the general result is also sufficient, and it allows for random arrivals of inspections.

Optimising over Incentive-Compatible Predictable Strategies

We establish the optimal predictable strategy using a recursive approach due to Davis (1993), where we solve for the optimal strategy for any given finite number of inspections, and then take the limit as this number grows large. We do this in two steps. First, we consider only histories at which the state is always in compliance. Second, we show how the principal's strategy can be adjusted to optimally respond to instances of non-compliance. Here, we provide a heuristic illustration of the first step, the formal arguments are contained in the proof of Theorem 1.

Between inspections, the evolution of promised utilities during continued periods of compliance are characterised by a pair of first-order differential equations. To see this, note that whenever there are no inspections, the principal's choice of dF_t cannot depend on the true state (conditional on the report $\hat{\theta}_t = 1$). Moreover, the principal and the agent are risk neutral, and therefore it is without loss to shift all fines into the future until after the next inspection, that is, $dF_t = 0$ for all $t \geq 0$ strictly before the next inspection. Assume additionally that $d\mu_t = 0$ whenever $dN_t^I = 0$. We later verify that the principal cannot lower monitoring costs through threats between inspections. Hence, if we start at $t = 0$ with initial values U_0^0 and U_0^1, the trajectories of the transitional payoff U_t^0 and the persistent payoff U_t^1 up until the first inspection are fully pinned down by the constraints (Pk) and (H) in Lemma 4.

This pair of coupled first-order differential equations has the following closed-form solution:

$$U_t^0 = e^{rt}(U_0^0 - \alpha(e^{\lambda t} - 1)(U_0^1 - U_0^0)) + c(e^{rt} - 1)/r, \tag{4.5}$$
$$U_t^1 = e^{rt}(U_0^1 + (1-\alpha)(e^{\lambda t} - 1)(U_0^1 - U_0^0)) + c(e^{rt} - 1)/r. \tag{4.6}$$

We now iterate over the number of inspections. Given the promised utilities are determined by (4.5) and (4.6), the problem reduces to the choice of initial values $(u^0, u^1) = (U_0^0, U_0^1)$ and inspection time T subject to the remaining constraints (O): $U_t^1 - U_t^0 \geq \frac{c}{\alpha\lambda}$ and (P): $U_t^i \in [-B, 0]$ for $i = 0, 1$. Note that full compliance is not achievable if the number of inspections is bounded so that the problem of minimising inspection costs subject to full compliance is ill-defined. To ensure that a solution exists for problem step $k \in \mathbb{N}$, (when the number of inspections cannot exceed k), we set the principal's objective to ensure compliance for as long as possible.

Consider the case with no inspection ($k = 0$). At any time $t > 0$, the principal has no possibility to distinguish the states, so that $U_t^0 = U_t^1$ and effort can never be incentivised. In consequence, if the principal has one inspection ($k = 1$), effort is achievable at most until the time of this inspection. The principal's goal is to perform this inspection as late as possible such that the payoff pair (U_t^0, U_t^1) fulfils conditions (O) and (P) up until this time. The trajectories in (4.5) and (4.6) can be combined to

$$U_t^1 - U_t^0 = (U_0^1 - U_0^0)e^{(r+\lambda)t},$$

showing that the obedience constraint (O) is fulfilled for all t if $u^1 - u^0 \geq \frac{c}{\alpha\lambda}$. Hence, for initial values (u^0, u^1), the optimal inspection time is given by the minimum of the two boundary hitting times

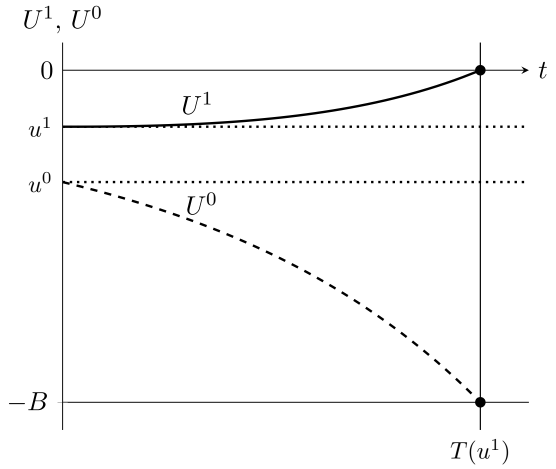
$$T^0(u^0, u^1) = \inf\{t > 0 : U_t^0 \leq -B\} \quad \text{and} \quad T^1(u^0, u^1) = \inf\{t > 0 : U_t^1 \geq 0\}.$$

Finally, the principal chooses (u^0, u^1) to maximise $\min\{T^0, T^1\}$. Note that, by (4.5), U_t^0 is increasing in u^0 while by (4.6), U_t^1 is decreasing in u^0, for all $t \geq 0$. An increase in u^0 increases both T^0 and T^1 so that it is optimal to set u^0 it as large as possible, i.e. $u^0 = u^1 - \frac{c}{\alpha\lambda}$.

The problem is thus reduced to finding the optimal initial utility u^1. Again, from (4.5) and (4.6) we observe that U_t^0 and U_t^1 are both increasing in u^1 for fixed value of $u^1 - u^0 = \frac{c}{\alpha\lambda}$. Thus, T^0 is increasing and T^1 decreasing in u^1. The minimum of the two hitting times, $T(u^1) \equiv \min\{T^0(u^1 - \frac{c}{\alpha\lambda}, u^1), T^1(u^1 - \frac{c}{\alpha\lambda}, u^1)\}$, is then maximised by choosing u^1 so that the utility paths hit the respective boundary simultaneously. Figure 4.1 illustrates this. For any other choice of initial value u_1 the hitting time is lower. For higher values u^1, the upper boundary is reached earlier, for lower values, the lower boundary is reached earlier.

Next, consider the case in which the principal can perform two inspections. If the first of two inspection time were chosen in the same way with both utility levels at the boundary, then U_t^1 would have to stay constant at 0 forever, no further incentives could be created despite another inspection being left. Hence, the principal-optimal initial value u^1 is strictly lower than in the previous case, so that the trajectory of U_t^0 reaches the lower boundary $-B$ before the trajectory of U_t^1 reaches the upper boundary 0. A lower level of u^1 forces the principal to inspect earlier, but she retains the option to fine the agent in the future, which is necessary for future inspections to be valuable.

Figure 4.1: The evolution of promised utilities over time conditional on continued compliance with a single inspection. Persistent utility is shown as solid line, transitional utility is dashed.



We now proceed in a similar fashion for further inspections, denoting by $u^1(k)$ the optimal initial value of the trajectory of U_t^1 when the total number of inspections is k. Inspection time T^k is determined by the time at which the trajectory of U_t^0 reaches $-B$. The more inspections are available to the principal, the more she will reduce the agents' persistent payoff in order to retain the option to fine him in the future. Iterating over the number of inspections k, we thus find that the optimal initial value $u^1(k)$ decreases as k increases. This implies that the inspection time T^k decreases. As k grows large, $u^1(k)$ converges to a unique limit u^*.

The optimal mechanism, conditional on continued compliance, has the property that the trajectory of the persistent utility is u-shaped, and it returns to the initial value u^* at the time of each inspection (see Figure 4.2). The limit values u^* and T^* are given by the solution to (4.5) and (4.6) with boundaries $(U_0^0, U_0^1) = (u^* - \frac{c}{\lambda\alpha}, u^*)$

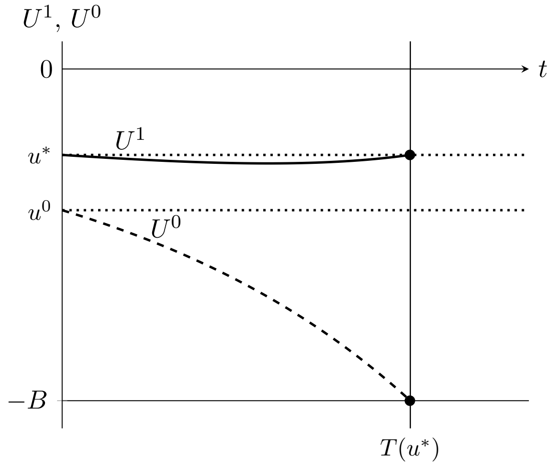
and $(U^0_{T^*}, U^1_{T^*}) = (-B, u^*)$:

$$U^0_t = e^{rt}u^* - e^{rt}(e^{\lambda t} - 1)\frac{1-\alpha}{\alpha}\frac{c}{\lambda} + (e^{rt} - 1)\frac{c}{r}, \tag{4.7}$$

$$U^1_t = e^{rt}u^* + e^{rt}(e^{\lambda t} - 1)\frac{1-\alpha}{\alpha}\frac{c}{\lambda} + (e^{rt} - 1)\frac{c}{r}. \tag{4.8}$$

We provide an implicit solution to T^* and u^* in Theorem 1 below. In case of a transition to non-compliance at time t, the promised utilities jump to $U^1_t = u^*$ and $U^0_t = u^* - \frac{c}{\lambda\alpha}$. Using continual fines and threats, the trajectories are held constant at these levels. In this way, upon another transition back to compliance, the promised utilities are already at their optimal initial values.

Figure 4.2: The evolution of promised utilities over time conditional on continued compliance with repeated inspections. Persistent utility is shown as solid line, transitional utility is dashed.



4.4 Characterisation of the Principal-Optimal Equilibrium

We now provide a characterisation of the principal-optimal equilibrium, and describe its properties and perform comparative statics. The equilibrium can be broadly summarised as follows. When the agent reports compliance, he does not have to pay a fine but is subject to periodic inspections. If the agent reports non-compliance, he pays a lump-sum fine that increases with proximity to the next inspection. While the agent is non-compliant, the principal requires the agent to pay a constant flow fine but conducts no inspections. The full details of the principal-optimal equilibrium are presented in the following theorem.

Theorem 1. *For H sufficiently large, there is a principal-optimal equilibrium, associated with period length $T^* > 0$ which is the unique positive solution to*

$$(Br - c)\,(1 - e^{-rT})\lambda\alpha - cre^{\lambda T}(e^{rT} - \alpha) + cr(1 - \alpha) = 0, \tag{4.9}$$

and utility level $u^* < 0$ given by

$$u^* = -B + 2e^{rT^*}(e^{\lambda T^*} - 1)\frac{1 - \alpha}{\alpha}\frac{c}{\lambda}v.$$

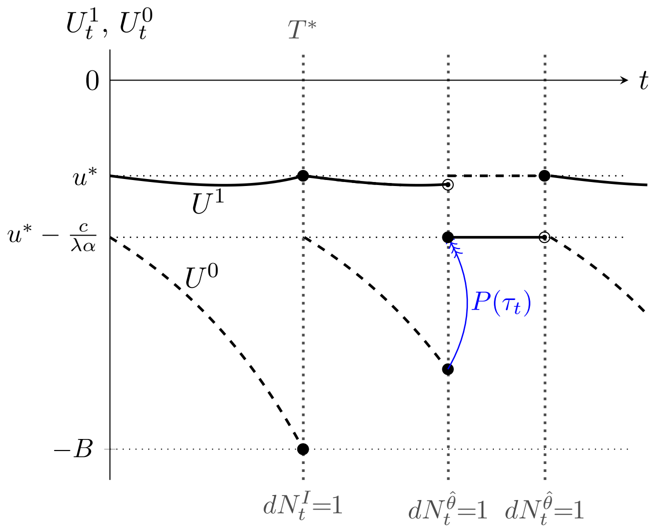
Let $\tau_t = t - \sup\{s \in (0, t) | \hat{\theta}_s = 0 \vee dN^I_s = 1\}$ be the time in compliance since the last transition or inspection. Along the equilibrium path, the following holds.

1. *If $\hat{\theta}_t = 1$: the agent pays no fine and an inspection is performed whenever $\tau_t = T^*$.*
2. *If $\hat{\theta}_t = 0$ and $dN^{\hat{\theta}}_t = 1$: the agent pays fine $P(\tau_t) = u^* - c/\lambda\alpha - U^0_{\tau_t}$ with $U^0_{\tau_t}$ given in (4.7).*
3. *If $\hat{\theta}_t = 0$ and $dN^{\hat{\theta}}_t = 0$: the agent pays a constant flow fine $r(c/\lambda\alpha - u^*)$.*

The implicit characterisation of the length of inspection cycles T^* and the initial level of promised utility u^* is obtained by combining (4.5) and (4.6). First note that the inspections are entirely predictable for the agent throughout. If the agent knows the time of the next inspection, he could of course try to hide a failure of compliance, in the hope of returning to compliance before the next inspection takes place. To deter the agent from such non-truthful behaviour, he faces a fine that increases over time.

Figure 4.3 depicts the evolution of the utilities for a specific path realisation. Starting in the good state, the agent makes no payments, an inspection occurs at time T^* (the first vertical line). At the second vertical line a decline from the good to bad state is reported. In a first step, the agent's utility drops to the current transitional payoff U^0_t. However, the agent pays fine $P(\tau_t)$ immediately so that his continuation utility increases by that amount to $u^* - \frac{c}{\lambda\alpha}$. The agent now pays a flow fine so that the promised utility levels remain constant at the optimal pair $(u^* - \frac{c}{\lambda\alpha}, u^*)$ until the next transition from bad to good state (the third vertical line). Conditional on the good state, the agent's expected utility is lowest in between two inspections. At the beginning, when the next inspection is still far in the future, the agent's incentives to delay the report of a decline in compliance is strongest so that the principal sets only moderate fines. As time progresses, the fine the agent would have to pay for a decline in compliance becomes more severe, decreasing his expected payoff. On the other hand, when the state is still good shortly before inspection,

Figure 4.3: The evolution of an example path realisation starting in the good state. Solid curves depict the persistent payoff, dashed curves depict the transitional payoffs



the agent is optimistic that a decline will not occur before the inspection and he will not be fined in this cycle.

4.4.1 Comparative Statics

We now consider the comparative statics of the principal-optimal equilibrium. We are primarily interested in the effects on inspection intensity and cost. Parameters affect the equilibrium in distinct ways. The payoff bound B and cost of effort c have a linear effect on payoffs and thus generate monotone effects on inspection intensity and cost. In contrast, adaptiveness α and variability λ determine the properties of the underlying stochastic process and thus produce more intricate effects.

Lemma 5. *Consider the equilibrium in Theorem 1 with length of inspection cycle T^*. Holding all other parameters fixed, T^* is*

- *increasing in B,*
- *decreasing in c,*
- *increasing in α, and*
- *increasing in λ for λ low enough and decreasing for high λ with $\lim_{\lambda\to\infty} T^*(\lambda) = 0$.*

The formal proof is relegated to the appendix and exploits the characterisation of T^* in Theorem 1 by implicit differentiation of (4.9). If the bound on the agent's disutility B increases, punishments become more effective. Unsurprisingly, the time between inspections T^* increases. The opposite happens when the flow cost c increases as effort becomes harder to incentivise. The comparative statics for α and λ show that the length of inspection cycles is increasing in the informativeness of inspections. When the effectiveness of effort α increases, each inspection becomes more informative about past effort choices, so that the time between inspections increases. In the case of the variability λ, the effect on the inspection cycle is non-monotone since the informativeness of inspections about past effort choice is itself non-monotone in λ. Informativeness is highest for intermediate levels of λ. When λ approaches zero, the state of compliance becomes extremely persistent and eventually becomes independent of effort. As λ grows large, the state of compliance becomes extremely fragile with a high frequency of transitions, so that inspections become increasingly uninformative about past effort choices.

In a manufacturing context, λ represents the pace of the industry and α represents the firm's internal adaptability to change. Short product cycles require dynamically evolving regulation. If this makes regulation less complex, compliance becomes easier to obtain. However, maintaining compliance for an extended period of time becomes harder.

Next, we study how the costs of compliance vary with changes in the parameters. Consider the expected discounted inspection costs in the case $\theta_0 = 1$[6] as a function of the equilibrium inspection time T^*:

$$
\begin{aligned}
C^H &= \mathbb{E}\left[\int_0^\infty e^{-rt}\kappa dN_t^I \mid \theta_0 = 1\right] \\
&= \frac{r+\lambda\alpha}{r(r+\lambda)} \cdot (r+\lambda(1-\alpha))\frac{e^{-(r+\lambda(1-\alpha))T^*}}{1-e^{-(r+\lambda(1-\alpha))T^*}} \cdot \kappa.
\end{aligned}
$$

The first fraction captures the relative likelihood of the good state. Future payoffs are effectively discounted at rate $r+\lambda(1-\alpha)$ to account for the possibility that the state may deteriorate prior to inspection. Intuitively, inspection costs decrease as T^* increases. For α and λ, there is an additional effect on costs as these parameters vary the stochastic process and thus the inspection costs caused by any fixed cycle length T.

Lemma 6. *Consider the equilibrium in Theorem 1 with inspection time T^*. Then the following holds for the discounted total inspection costs C^H.*

- *C^H is decreasing in B.*
- *C^H is increasing in c.*
- *There exist $\underline{\alpha} > \frac{c(r+\lambda)}{Br\lambda}$ and $\bar{\alpha} \in [\underline{\alpha}, 1]$ such that C^H decreases in α if $\alpha < \underline{\alpha}$ and increases if $\alpha > \bar{\alpha}$.*

[6] The case $\theta_0 = 0$ is analogous. We have $C^L = \frac{\lambda\alpha}{r+\lambda\alpha}C^H$.

- *There exists* $\underline{\lambda} > \frac{cr}{Br\alpha - c}$ *such that* C^H *decreases in* λ *if* $\lambda < \underline{\lambda}$ *and* C^H *goes to infinity for* $\lambda \to \infty$.

The formal proof is relegated to the appendix. The fractions in the third and fourth statement are rearrangements of our standing feasibility assumption $B > \frac{c(r+\lambda)}{\alpha r \lambda}$. In the case of B and c, the results follow immediately from Lemma 5. For the effectiveness of effort α, there are several effects on inspection costs that work in opposite directions. As α increases, the agent remains in compliance for longer periods of time, and more easily regains compliance. Since inspections are performed only when the agent reports compliance, inspection costs increase if we keep the cycle length T fixed. On the other hand, by Lemma 5, the inspection cycle length grows in α which decreases inspection costs. The third bullet in Lemma 6 states that the second effect dominates for low values of α and the first effect for high values, so that the effect of α on inspection costs is overall non-monotone.

The variability of compliance λ affects inspection costs directly through transitions in compliance and indirectly through the length of inspection cycles T^*. The direct effect of an increase in λ is a reduction in inspection costs, because inspection cycles are increasingly often interrupted so that inspections are less likely to be carried out. For low variability, an increase in λ results in a longer equilibrium inspection cycle (for low λ), and thus the total effect of an increase in λ on inspection costs must be negative. For high variability, an increase in λ causes the cycle length to decrease, so that the two effects go in opposite directions. The last bullet in the lemma shows that the cost-increasing effect indeed dominates for large values of λ. Note that for fixed $T > 0$ the total costs C^H decrease to 0 as λ grows arbitrarily large. Lemma 6 shows that T^* approaches zero fast enough so that the inspection costs explode in the limit.

This cost increase arises because inspections are scheduled periodically so that the agent has positive time to aim for a transition to compliance prior to inspection. This deviation is more attractive for high values of λ. As λ becomes arbitrarily high, the probability of reaching any time in the strict future without a prior state change vanishes. For inspections to happen with high enough probability to incentivise effort, time T^* has to shrink to 0. The intuition how this leads to arbitrarily large inspection costs is the following. For the agent who considers to shirk for an instant, the payoff effect at the next inspection is effectively discounted at rate $r+\lambda$. Current effort affects the state at inspection only if no further transition occurs previously. This implies that the limit of $\lambda e^{-(r+\lambda)T^*(\lambda)}$ converges to a finite constant, as the proof shows. However, the effective discount rate on path – i.e. the relevant rate to evaluate the principal's costs – is only $r + \lambda(1-\alpha)$ so that the limit of C^H as λ goes to ∞ is proportional to $\lim \lambda e^{-(r+(1-\alpha)\lambda)T^*(\lambda)}$ which is infinity for $\alpha < 1$ since $\lambda e^{-(r+\lambda)T^*(\lambda)}$ converges.

As the comparative statics uncovered, the high costs of compliance in the case of high λ stem from the agent's opportunity to achieve a good outcome with high probability whenever the next inspection is not imminent. Making inspections imminent (choosing T^* close to 0) makes the costs explode. This suggests that randomisation over inspection times may be valuable as it permits to threaten the possibility of instant inspection at all times without having to perform inspections at all times.

4.5 Overcoming the Commitment Problem

By Lemma 1, choosing an inspection schedule which is non-predictable for the agent cannot decrease the principal's costs when she lacks commitment power. This section shows that, with commitment power, there is value in non-predictable inspections. Compliance can be attained with fewer inspections on average. Before we turn to the analysis, we discuss possible sources of sufficient commitment power for randomisation to be feasible for the principal and situations in which we should expect them.

The proof of Lemma 1 uncovers that profitable randomisation is unsustainable in equilibrium because deviations by the principal become unobservable. This suggests a public randomisation device to force the principal to follow a random inspection schedule where agent and principal agree on the set of realisations that lead to inspection. The principal can be punished with immediate contract termination if he does not comply. The use of (public) randomisation devices is common in the theoretical literature on repeated interactions. In practice, such methods are rarely observed. Our analysis suggests a reason in the context of compliance: benefiting from random inspections requires a randomisation device that is perfectly observed by both parties – there cannot be any idiosyncratic noise as the possibility of disagreement regarding the realisation necessarily leads to non-compliance on path. Yet, our analysis suggests another possibility to profit from random inspections. Lemma 1 does *not* rule out randomisation altogether. Rather, it shows that the required indifference between conducting the inspection and postponing it precludes the principal from benefiting from randomisation. The separation of inspection planning and its execution can achieve this. For example, inspections can be prescribed by the internal compliance manager but carried out by external practitioners.[7] This separates the inspection cost from the inspection decision and, thus, eliminates the principal's incentive to skip inspections.

There is an important difference to note between this suggestion and two seemingly analogous alternatives: directly compensating the inspection cost to the prin-

[7]See European Commission (2019, p. 24) for this suggestion in the context of ICPs for dual use goods manufacturers.

cipal at one extreme and outsourcing the oversight activity at the other. Compensating the principal for performed inspections resolves the problem highlighted above theoretically. However, this would require precise knowledge of the cost and effort required by the principal to perform an inspection (cost κ in our model). If the compensation for an inspection falls below this value, the incentive to skip it remains. If the compensation is too generous, this creates an incentive to inspect inefficiently often. An example of the separation of planning and execution can be found in the context of banking supervision in Germany. Depending on the bank's size, the European Central Bank or the supervisory agency at the federal Finance Ministry (BaFin) fulfils the supervisory function and is responsible for scheduling audits. The execution of these audits, however, is always done by the German Bundesbank (BaFin, 2016).

Next we provide a random mechanism and show that the principal can decrease the inspection costs required to implement compliance in comparison to predictable rules. The derivation remains agnostic about the possible source of commitment power that allows for randomisation. Assuming that the principal has full commitment power over random inspections, consider the stationary stochastic mechanism in which the principal sets no fines and no threats between inspections, and chooses the stationary inspection level m^* such that the agent's promised utility remains constant at $(\bar{U}^0, \bar{U}^1)$. Combining the promise-keeping conditions with the honesty and obedience constraints under the assumption of equality in each of them, we find that a stationary mechanism must solve

$$\begin{aligned}
0 &= r\bar{U}^1 + \lambda(1-\alpha)(\bar{U}^1 - \bar{U}^0) + c, \\
0 &= r\bar{U}^0 - \lambda\alpha(\bar{U}^1 - \bar{U}^0) + m^*(B + \bar{U}^0) + c \\
0 &= \lambda\alpha(\bar{U}^1 - \bar{U}^0) - c.
\end{aligned}$$

A simple calculation reveals that this system of equations has the solution

$$\bar{U}^1 = -\frac{c}{r\alpha}\ , \quad \bar{U}^0 = -\frac{c}{r\alpha} - \frac{c}{\lambda\alpha}\ , \quad m^* = \frac{cr(r+\lambda)}{Br\lambda\alpha - c(r+\lambda)}. \tag{4.10}$$

In the stationary mechanism characterised by these three equations, the honesty and the obedience conditions bind at all t. The stochastic scheme is structurally similar to the predictable rule in the sense that the agent is fined for non-compliance only, and inspections are performed when he reports compliance. The stochastic mechanism is fully stationary, however. Inspections are conducted at a constant rate in such a way that promised utilities remain constant throughout.

The value of random inspections can be observed most strikingly when we consider the limit as λ grows large. According to Lemma 6, the inspection costs grow without bound in the predictable equilibrium. With the stochastic mechanism, the inspection costs starting in the compliant state are given by

$$\frac{r+\lambda\alpha}{r(r+\lambda)}m^*\kappa = \frac{c(r+\lambda\alpha)}{Br\lambda\alpha - c(r+\lambda)}\kappa.$$

This term is always decreasing in λ, converging to $\frac{c\alpha}{Br\alpha - c}\kappa$. As explained with the comparative statics of the predictable equilibrium, this contrast stems from the impossibility to threaten immediate inspections without unbounded costs when randomisation is infeasible.

4.6 Concluding Remarks

The paper studies incentive schemes in a dynamic principal-agent setting with costly monitoring in which the principal wants to induce maximal compliance at minimal cost. The principal provides incentives for effort by performing inspections and setting fines. We show that full compliance is attainable even without commitment power. The cost-minimal full-compliance equilibrium is derived by solving an auxiliary mechanism-design problem in which the principal is restricted to using predictable inspection rules. Without commitment power, the principal cannot profit from randomised inspections even though this is profitable with commitment power. Our findings identify a possible reason for the failure of effective enforcement: the lack or wrong assessment of commitment power paired with the use of random inspection schedules. As predictable schedules make deviations by the regulator easier to observe, their commitment requirements are significantly weaker.

A crucial assumption maintained throughout is that the principal seeks full compliance, defined as maximum effort and truthful reporting after any history along the equilibrium path. This assumption keeps the analysis tractable. We interpret the results in this paper as a benchmark for the design of optimal policies, and as a sound theoretical approach to generating predictions of the costs of effective enforcement. For a subset of the parameter space, however, full compliance would not be socially optimal. Extending the analysis to allow for periods of non-compliance by the agent would be interesting, but it makes the underlying optimisation substantially more complicated.

There are other aspects not considered in this paper. For example, the analysis could be extended to allow for exogenous signals about compliance, for an agent who is imperfectly informed, or an imperfect monitoring technology. Some of these issues are discussed in Varas et al. (2020). These and other variants might be fruitful avenues for future research for which the characterisation of incentive-compatible contracts provided in this paper may result useful.

4.A Appendix

4.A.1 Strategies and outcomes

This part of the appendix contains the formal restrictions on the players' strategy spaces to ensure that any combination of strategies leads to a unique and well-defined outcome. We adopt the approach by Kamada and Rao (2018), requiring that actions are not changed 'too frequently' on any time interval. To apply this approach we first need to restrict the continuous action of each player, fines and effort.

We say that for a history $h_t \in H_t$, there is an *intervention for the agent* if either $t = 0$ or if $t > 0$ and at least one of the following holds: (i) $\theta_t - \theta_{t-} \neq 0$, (ii) $\hat{\theta} - \hat{\theta}_{t-} \neq 0$, (iii) $N_t^I - N_{t-}^I \neq 0$. Similarly, we say that there is an *intervention for the principal* if $t = 0$ or if $t > 0$ and at least one of the properties (ii) and (iii) holds.

For the principal, no new information arrives in between interventions, and we restrict the strategy of fines to reflect this by being predictable in between interventions. Formally, for any two histories h_t, h_t': $f_t(h_t) \neq f_t(h_t')$ only if there exists $\tau \leq t$ such that τ is an in an intervention time for the principal and the truncation of the above histories at time τ, h_τ and h_τ', are distinguishable for the principal. Put differently, the principal's fines are specified pathwise; at each intervention, she chooses how the fines proceed until the time of the next intervention. Given that no information arrives between interventions any optimally chosen path at t will be sequentially optimal until the next intervention. Similarly for the agent, we restrict the effort strategy to be predictable in between interventions: For any two histories h_{t-}, h_{t-}': $e_t(h_{t-}) \neq e_t(h_{t-}')$ only if there exists $\tau < t$ such that τ is an in an intervention time for the agent and $h_\tau \neq h_\tau'$.

To adapt the approach by Kamada and Rao (2018), we require strategies to fulfil the properties *traceability* and *frictionality* defined below. Lemma 7 shows that any combination of strategies that fulfils these restrictions yields a well-defined and unique outcome path.

A history h is said to be *consistent* with the agent's strategy (r, e) at time t if $r_t(h_{t-}, \theta_t) = \hat{\theta}_t$ and $e_t(h_t) = \eta_t$. Similarly, a history h is consistent with the principal's strategy (n, f) at time t if $n_t(\pi, h_{t-}, \hat{\theta}_t) = dN_t^I$ and $f_t(h_t) = dF_t$.

Definition 1. *The agent's strategy (r, e) is* ***traceable*** *if for any time-t history h_t and any principal-action path $\{N_s^I, F_s\}_{s\geq 0}$ that coincides with h_t for all $s < t$, there is a continuation path $\{\hat{\theta}_s, \eta_s\}_{s\geq t}$ that is consistent with (r, e). Analogously, The principal's strategy (n, f) is traceable if for any realisation of π, any time-t history h_t, and any agent-action path $\{\hat{\theta}_s, \eta_s\}_{s\geq 0}$ that coincides with h_t for all $s < t$: there is a continuation path $\{N_s^I, F_s\}_{s\geq t}$ that is consistent with (n, f).*

Definition 2. *The agent's strategy (r, e) is* ***frictional*** *if for any time-t history h_t, there is conditional probability one that the report path $\{\hat{\theta}_s\}_{s\geq t}$ has only finitely many report changes on any finite interval $[t, u]$ for all paths $\{\eta_s, \hat{\theta}_s\}_{s\geq t}$ such that there is a principal-action path $\{N_s^I, F_s\}_{s\geq t}$ for which the history $\left(h_{t-}, \{N_s^I, F_s\}_{s\geq t}, \{\eta_s, \hat{\theta}_s\}_{s\geq t}\right)$ is consistent with the agent's strategy. Analogously, the principal's strategy (n, f) is frictional if for any time-t history h_t, there is conditional probability one that the inspection path $\{N_s^I\}_{s\geq t}$ has only finitely many inspections on any finite interval $[t, u]$ for all paths $\{N_s^I, F_s\}_{s\geq t}$ such that there is an agent-action path $\{\eta_s, \hat{\theta}_s\}_{s\geq t}$ for which the history $\left(\pi, h_{t-}, \{N_s^I, F_s\}_{s\geq t}, \{\eta_s, \hat{\theta}_s\}_{s\geq t}\right)$ is consistent with the principal's strategy.*

Lemma 7 (Existence and Uniqueness of consistent outcome paths)**.** *Given any possible history $h_{u-} = \left\{\pi, z_t, \eta_t, \hat{\theta}_t, N_t^I, F_t\right\}_{t\in[0,u)} \cup \{\eta_u\}$, any combination of strategies $((r, e), (n, f))$ that are traceable and frictional yields a unique consistent path $\left(\{\eta_t\}_{t\in(u,\infty)}, \{\hat{\theta}_t, N_t^I, F_t, \}_{t\in[u,\infty)}\right)$ almost surely.*

Proof. **Step 1: Uniqueness.** Fix a pair of strategies, a history up to u, and any realisation of the shock process $\{z_t\}_{t\in[u,\infty)}$. Suppose there are two distinct continuation paths $x =$

$\{\eta_t^x, \hat{\theta}_t^x, N_t^{I^x}, F_t^x\}_{t\in[u,\infty)}$ and $y = \{\eta_t^y, \hat{\theta}_t^y, N_t^{I^y}, F_t^y\}_{t\in[u,\infty)}$ that are consistent with the strategies and the shock path. Let $\underline{t} = \inf\{t \geq u : x_t \neq y_t\}$ be the first time at which the processes differ. Strategy e maps history $h^A_{t^A_k}$ into a deterministic process $\{\eta_s\}_{s\in(t^A_k,\infty)}$ only for times t^A_k at which an intervention for the agent occurs. Likewise, strategy f maps history $h_{t^P_k}$ into a deterministic process $\{F_s\}_{s\in[t^P_k,\infty)}$ for times t^P_k with an intervention for the principal. Therefore, if $\eta^x_s \neq \eta^y_s$ for $s > u$ or $F^x_s \neq F^y_s$ for $s \geq u$, then there must also be a time $t \leq s$ with an intervention at t, i.e. $\exists k \in \mathbb{N}$ s.t. $t = t^A_k$ or $t = t^P_k$. Furthermore, we must have $h^x_t \neq h^y_t$ at this intervention. With probability 1, the realisation $\{z_t\}_{t\in[u,\infty)}$ has only finitely many jumps on any closed interval. Hence, by frictionality, there are at most finitely many interventions on any closed interval. Therefore, $\underline{t}$ defined above must be an intervention time and the infimum is attained, i.e. $x_{\underline{t}} \neq y_{\underline{t}}$. We therefore must have $\hat{\theta}^x_{\underline{t}} \neq \hat{\theta}^y_{\underline{t}}$ or $N^{I^x}_{\underline{t}} \neq N^{I^y}_{\underline{t}}$ and, as $\underline{t}$ is the first such time, $h^x_{\underline{t}-} = h^y_{\underline{t}-}$. As $\hat{\theta}^x_{\underline{t}}$ and $\hat{\theta}^y_{\underline{t}}$ both result from the same strategy, this, however, implies that $\hat{\theta}^x_{\underline{t}} = \hat{\theta}^y_{\underline{t}}$, leaving as only possibility that $N^{I^x}_{\underline{t}} \neq N^{I^y}_{\underline{t}}$. This contradicts consistency of both processes with the fixed strategy (as $h^x_{\underline{t}-} = h^y_{\underline{t}-}$). Hence, any pair of traceable and frictional strategies gives at most one consistent outcome.

Step 2: Existence. Existence of a consistent outcome path is shown constructively: Start with arbitrary history $h_{u-} = \left\{\pi, z_t, \eta_t, \hat{\theta}_t, N^I_t, F_t\right\}_{t\in[0,u)} \cup \{\eta_u\}$ and fix a realisation of the shock process $\{z_t\}_{t\in[u,\infty)}$. We apply the steps below iteratively until they give an outcome path consistent with z and the strategies for $t \geq u$: Define paths $\{\eta^0_t, \hat{\theta}^0_t, N^{I^0}_t, F^0_t\}$ equal to the history up to u and such that for $t > 0u$: $\eta^0_t = e_t(h_{\max_k t^A_k < u})$, and for $t \geq u$: $\hat{\theta}^0_t = \hat{\theta}_{u-}$, $N^{I^0}_t = N^I_{u-}$ and $dF^0_t = f_t(h_{\max_k t^A_k < u})$.[8]

Let $n = 1$ and t(1) = u.

i) By traceability, there are paths $\{\eta^n_t, \hat{\theta}^n_t\}_{t\geq 0}$ such that, for $t < t(n)$: $\{\hat{\theta}^n_t, \eta^n_t\} = \{\eta^{n-1}_t, \hat{\theta}^{n-1}_t\}$ and that $\{\eta^n_t, \hat{\theta}^n_t, N^{I^{n-1}}_t, F^{n-1}_t\}_{t\geq 0}$ is consistent with the agent's strategy and process z for $t \geq t(n)$. Set $\{\eta^n_t, \hat{\theta}^n_t\}$ equal to these processes.

   Similarly, traceability implies that there exist paths $\{N^{I^n}_t, F^n_t\}$ with $(N^{I^n}_t, F^n_t) = (N^{I^{n-1}}_t, F^{n-1}_t)$ for $t < t(n)$ and such that $\{\eta^n_t, \hat{\theta}^n_t, N^{I^n}_t, F^n_t\}_{t\geq 0}$ is consistent with the principal's strategy on $t \geq u$. Set $\{N^{I^n}_t, F^n_t\}$ equal to these processes and continue to step (ii).

ii) If $\{\eta^n_t, \hat{\theta}^n_t, N^{I^n}_t, F^n_t\}$ is consistent with the strategies for all $t \in [u,\infty)$, stop the procedure. The proof is complete. Otherwise, redefine $n = n + 1$ and set $t(n + 1)$ equal to the largest time v such that there is an intervention at v and $\{\eta^n_t, \hat{\theta}^n_t, N^{I^n}_t, F^n_t\}$ is consistent with the strategies for all $t \in [u, v)$, go to step (i).

If the above procedure stops after finite n, that is because of having given a consistent process and the proof is complete. In the case in which it does not stop after finitely many iterations, $\lim_{n\to\infty}\{\eta^n_t, \hat{\theta}^n_t, N^{I^n}_t, F^n_t\}_{t\geq 0}$ is consistent with the strategies on $[u,\infty)$ with probability one. To see this, note that for every n, $t(n+2) > t(n)$. Given that, with probability one, any finite interval has only finitely many interventions, $\lim_{n\to\infty} t(n) = \infty$ which implies consistency of the resulting process for all $t \in [u,\infty)$. □

4.A.2 Proofs from main part

Proof of Lemma 1

Suppose there exists a perfect Bayesian equilibrium $\{\tilde{n}_t, \tilde{f}_t, \tilde{e}_t, \tilde{r}_t\}_{t\geq 0}$ – possibly with random inspections – that induces full compliance and generates a higher payoff. Let $\tilde{\nu}^I_t$ be the compensator

[8] That is, report and inspections are held constant from u onward and fines and effort are chosen according to the strategies (depending only on the last intervention before u) for the case that no further interventions occur.

associated with process $\tilde{N}_t^I$. We can find a process $\hat{N}^I$ with the property that for all $s \geq t$,

$$\hat{N}_s^I - \hat{N}_t^I = 0 \Longleftrightarrow \tilde{\nu}_s^I - \tilde{\nu}_t^I = 0.$$

Intuitively, for an equilibrium with random inspections, $\hat{N}^I$ is the path realisation of inspections with the smallest gaps. Because $\hat{N}^I$ is the realisation of an equilibrium strategy, the principal's payoff from using $\hat{N}^I$ is the same as in the equilibrium. Moreover, $\hat{N}_t^I$ is predictable by construction. The agent is inspected more than at any other path realisation so that this strategy induces compliance since $\{\tilde{N}_t^I, \tilde{F}_t, \tilde{\eta}_t, \hat{\theta}_t\}_{t\geq 0}$ was part of a maximal compliance equilibrium. □

Proof of Lemma 2

We show that any predictable incentive-compatible principal-strategy that generates a positive value for the principal at each t can be implemented in equilibrium. Let $\{N_t, F_t\}$ be the paths induced by the strategy in the lemma. By hypothesis (ii), compliance is incentive compatible for the agent. Let $\tilde{n}$ be an alternative inspection strategy for the principal, possibly random, and let $\tilde{N}^I$ be the resulting inspection path if the agent follows the compliant strategy. The set $D = \{t | dN_t^I \neq d\tilde{N}_t^I\}$ represents the dates at which the principal observably deviates. Adapt the agent's strategy so that $\eta_t = 1$ at all $t \leq \inf D$ and $\eta_t = 0$ otherwise. Since the payoff for the principal from the strategy in the lemma is positive at each t, and the payoff from any deviating strategy is equal for all $t < \inf D$, her deviation cannot be profitable as it results in a payoff of 0 from $\geq \inf D$ onward. Finally, adapt the principal's strategy from the lemma such that he fines the agent as harshly as possible whenever the agent was expected to exit but failed to do so. This way, for the agent the strategy with $\eta_t = 0$ for $t > \inf D$ is incentive compatible and the constructed equilibrium differs from the initial strategy profile in Lemma 2 at most off the equilibrium path. □

Proof of Lemma 3

Denote by $\mathcal{F}$ the filtration generated by the random processes θ, $\hat{\theta}$ and ν^I. Define

$$W_t := \int_0^t e^{-rs}(-dF_s - c\eta_s ds) + e^{-rt}U_t.$$

The corresponding representation in differential form is

$$dW_t = e^{-rt}(-dF_t - c\eta_t dt) - re^{-rt}U_t + e^{-rt}dU_t. \tag{4.A.1}$$

The process $\{W_t\}$ is an $\mathcal{F}$-martingale by construction. By the martingale representation theorem for marked point processes (Last and Brandt, 1995, Theorem 1.13.2), there exist $\mathcal{F}$-predictable functions $\tilde{\Delta}_t^\theta$, $\tilde{\Delta}_t^{\hat{\theta}}$ and $\tilde{\Delta}_t^I$ such that

$$dW_t = \sum_{a\in\{\theta,\hat{\theta},I\}} \tilde{\Delta}_t^a(dN_t^a - d\nu_t^a) \tag{4.A.2}$$

Replacing $\tilde{\Delta}_t^a = e^{-rt}\Delta_t^a$ and then equating (4.A.1) and (4.A.2) yields

$$dU_t = rU_t dt + dF_t + c\eta_t dt + \sum_{a\in\{\theta,\hat{\theta},I\}} \Delta_t^a(dN_t^a - d\nu_t^a).$$

This is the representation of the evolution of promised utilities shown in the lemma. □

Proof of Lemma 4

The following lemma provides an intermediate step toward the proof of Lemma 4:

Lemma 8. *A mechanism that induces the payoffs $\{U_t\}_{t\geq 0}$ is incentive compatible with maximum effort and truthful reporting if and only if for all $t \geq 0$:*

(i) $(r+q^*)\Delta_t^{\hat{\theta}} - d\nu_t^I(\Delta_t^I - \Delta_t^{\hat{\theta}}) \geq d\Delta_t^{\hat{\theta}}$ *when* $\theta_t \neq \hat{\theta}_t$,

(ii) $(1-2\theta_{t-})\lambda\alpha(\Delta_t^{\theta} + \Delta_t^{\hat{\theta}}) \geq c$ *when* $\theta_t = \hat{\theta}_t$,

(iii) $U_t \in [-B, 0]$.

Proof. Define

$$W_t = \int_0^t e^{-rs}(-dF_s - c\eta_s ds) + e^{-rt}\tilde{U}_t.$$

to be the agent's expected payoff from using effort strategy $\{\tilde{\eta}_s\}$ and reporting strategy $\{\hat{\theta}_s\}$ up to time t with maximum effort and truthful reporting thereafter. Here $\tilde{U}_t$ is the expected continuation payoff. We may have $\tilde{U}_t \neq U_t$ if the agent has reported non-truthfully, i.e. $\hat{\theta}_{t-} \neq \theta_{t-}$.

Consider first the case in which the agent's report regarding his type at time t is truthful, so that $\tilde{U}_t = U_t$. Differentiating with respect to t yields

$$dW_t = e^{-rt}(-dF_t - c\eta_t dt) - re^{-rt}U_t dt + e^{-rt}dU_t.$$

Using Lemma 3 to replace dU_t yields

$$\begin{aligned} dW_t &= \left(e^{-rt}(-dF_t - c\eta_t dt) - re^{-rt}U_t dt + e^{-rt}\left(rU_t dt + dF_t + cdt + \sum_{a\in\{\theta,\hat{\theta},I\}} \Delta_t^a(dN_t^a - d\nu_t^a)\right)\right) \\ &= e^{-rt}\left((1-\eta_t)cdt + \sum_{a\in\{\theta,\hat{\theta}\}} \Delta_t^a(dN_t^a - q_t^* dt) + \Delta_t^I(dN_t^I - d\nu_t^I)\right), \end{aligned}$$

where $q_t^* = q_t(1)$. If the agent deviates for an additional instant (but still reports truthfully) then

$$dN_t^{\theta} = dN_t^{\hat{\theta}} = \begin{cases} 1 & \text{with probability } q_t(\tilde{\eta}_t)dt \\ 0 & \text{with probability } 1 - q_t(\tilde{\eta}_t)dt \end{cases}.$$

Taking expectations therefore yields

$$\mathbb{E}_t[dW_t] = e^{-rt}\mathbb{E}\left[(1-\eta_t)cdt + (\Delta_t^{\theta} + \Delta_t^{\hat{\theta}})(q_t(\tilde{\eta}_t) - q_t^*)dt\right].$$

It follows from Condition (ii) that

$$(\Delta_t^{\theta} + \Delta_t^{\hat{\theta}})q(\tilde{\eta}_t) - c\eta_t \leq (\Delta_t^{\theta} + \Delta_t^{\hat{\theta}})q_t^* - c.$$

Thus $\mathbb{E}_t[dW_t] \leq 0$. We thus obtain the chain of inequalities

$$\mathbb{E}_0[W_t] = \mathbb{E}_0\left[\int_0^t dW_s + W_0\right] = \int_0^t \mathbb{E}_0[dW_s] + \mathbb{E}_0[W_0] = \int_0^t \mathbb{E}_0\Big[\mathbb{E}_s[dW_s]\Big] + W_0 \leq W_0. \quad (4.A.3)$$

Now, consider the case in which the agent's most recent report about quality at time t is false, that is $\theta_{t-} \neq \hat{\theta}_{t-}$ and he continues the non-truthful strategy for an additional moment at time t. If no change in quality occurs at the additional moment, then the agent must correct his report immediately thereafter. If a change in quality occurs, then the previously false statement becomes truthful, and thus his report does not change. Therefore, we have the following:

$$\begin{aligned} d\tilde{U}_t &= \tilde{U}_t - \tilde{U}_{t-dt} \\ &= dN_t^\theta(U_t - U_{t-dt} - \Delta_{t-dt}^{\hat{\theta}}) + dN_t^I(U_t + \Delta_t^I - U_{t-dt} - \Delta_{t-dt}^{\hat{\theta}}) \\ &\qquad + (1 - dN_t^\theta - dN_t^I)(U_t + \Delta_t^{\hat{\theta}} - U_{t-dt} - \Delta_{t-dt}^{\hat{\theta}}) \\ &= dN_t^\theta(dU_t + d\Delta_t^{\hat{\theta}} - \Delta_t^{\hat{\theta}}) + dN_t^I(dU_t + d\Delta_t^{\hat{\theta}} - \Delta_t^{\hat{\theta}} + \Delta_t^I) + (1 - dN_t^\theta - dN_t^I)(dU_t + d\Delta_t^{\hat{\theta}}) \\ &= dU_t + d\Delta_t^{\hat{\theta}} - dN_t^\theta \Delta_t^{\hat{\theta}} + dN_t^I(\Delta_t^I - \Delta_t^{\hat{\theta}}). \end{aligned} \tag{4.A.4}$$

Using again Lemma 3 to replace dU_t, we obtain

$$\begin{aligned} dW_t &= e^{-rt}(-dF_t - c\eta_t dt) - re^{-rt}(U_t + \Delta_t^{\hat{\theta}}) \\ &\qquad + e^{-rt}\Big(rU_t dt + dF_t + cdt + \Delta_t^\theta(dN_t^\theta - q_t^*) + d\Delta_t^{\hat{\theta}} - dN_t^\theta \Delta_t^{\hat{\theta}} + dN_t^I(\Delta_t^I - \Delta_t^{\hat{\theta}})\Big) \end{aligned}$$

It follows from the honesty constraint (i) that, in expectation, $d\Delta_t^{\hat{\theta}} \le (r + q^*)\Delta_t^{\hat{\theta}} - d\nu_t(\Delta_t^I - \Delta_t^{\hat{\theta}})$. When we substitute it into dW_t and simplify, using again $\tilde{U}_t = U_t + \Delta_t^{\hat{\theta}}$, we obtain

$$\mathbb{E}_t[dW_t] = e^{-rt}\left((1-\eta)cdt + (\Delta_t^\theta - \Delta_t^{\hat{\theta}})q(\tilde{\eta}_t) - q^*(\Delta_t^\theta - \Delta_t^{\hat{\theta}})\right).$$

Now, $\Delta_t^\theta - \Delta_t^{\hat{\theta}} = (\Delta_t^\theta + U_t) - (\Delta_t^{\hat{\theta}} + U_t)$ is the payoff difference from a change in quality without a change in report and a change in report without a change in quality. Since $\theta_{t-} \neq \hat{\theta}_{t-}$ by hypothesis, this is identical to $\tilde{\Delta}_t^\theta + \tilde{\Delta}_t^{\hat{\theta}}$ after the history in which the agent's true quality was identical to his reported quality. Thus (ii) implies that $\eta_t = 1$ maximises the right-hand side, so that $\mathbb{E}_t[dW_t] \le 0$. By the same argument as in (4.A.3), we have

$$\mathbb{E}_0[W_t] \le W_0 = U_0.$$

so that the agent cannot profit from deviating. Taking the limit, we find that

$$\lim_{t\to\infty} \mathbb{E}_0[W_t] \le U_0.$$

which implies that the agent cannot gain from deviating from maximum effort and truthful reporting. Conversely, if the incentive constraint (i) is violated, then the above inequalities are inverted, so that the agent has a strict incentive to be dishonest. □

To conclude the proof of Lemma 4, we show that condition (Pk) follows from Lemma 3 and (H), (H) and (P) are equivalent to conditions (i), (ii) and (iii) in Lemma 8. Consider a contract and a strategy for the agent that jointly generate the payoff process $\{U_t\}_{t\ge0}$ for the agent, and denote by $\{U_t^1, U_t^0\}_{t\ge0}$ the associated pair of promised utilities defined in Equation (4.4).

(1.) By the definition of U_t^0, U_t^1, we have

$$\Delta_t^\theta + \Delta_t^{\hat{\theta}} = \begin{cases} U_t^1 - U_t^0 & \text{if } \theta_{t-} = \hat{\theta}_{t-} = 0 \\ U_t^0 - U_t^1 & \text{if } \theta_{t-} = \hat{\theta}_{t-} = 1 \end{cases}, \qquad q_t^* = q_t(1) = \begin{cases} \alpha\lambda & \text{if } \theta_{t-} = 0 \\ (1-\alpha)\lambda & \text{if } \theta_{t-} = 1 \end{cases}. \tag{4.A.5}$$

Combining these two expressions, we can write more succinctly:

$$q_t^*(\Delta_t^\theta + \Delta_t^{\hat{\theta}}) = \lambda(\theta_{t-} - \alpha)(U_t^1 - U_t^0).$$

Lemma 3 implies that, conditional on the event that $dN_t^\theta = dN_t^{\hat\theta} = dN_t^I = 0$, we have

$$\begin{aligned} dU_t^i &= rU_t^i dt - q_t^*(\Delta_t^\theta + \Delta_t^{\hat\theta}) + dF_t + cdt - d\nu_t \Delta_t^I \\ &= rU_t^i dt + \lambda(i-\alpha)(U_t^1 - U_t^0) + dF_t + cdt - d\nu_t \Delta_t^I \end{aligned}$$

which proves condition (Pk) in Lemma 4. Note that Δ_t^I measures the difference in utility before and after an inspection when the agent reports his type truthfully.

(2.) Next, suppose that the agent is not truthful after some history at time t. Let $i = \theta_t$ be true quality and suppose the agent reports $j = 1 - \theta_t$. Then, $U_t^i = U_t + \Delta_t^{\hat\theta}$, and

$$\begin{aligned} dU_t^i &= (U_{t+dt} + \Delta_{t+dt}^{\hat\theta}) - (U_t + \Delta_t^{\hat\theta}) \\ &= rU_t dt + dF_t + cdt - q_t^* \Delta_t^\theta + d\Delta_t^{\hat\theta} \\ &\leq rU_t dt + dF_t + cdt - q_t^* \Delta_t^\theta + (r + q^*)\Delta_t^{\hat\theta} - d\nu_t(\Delta_t^I - \Delta_t^{\hat\theta}) \qquad (4.A.6) \\ &= r(U_t + \Delta_t^{\hat\theta}) - q_t^*(\Delta_t^\theta - \Delta_t^{\hat\theta}) - d\nu_t(\Delta_t^I - \Delta_t^{\hat\theta}) + dF_t + cdt \\ &= rU_t^i + \lambda(i-\alpha)(U_t^1 - U_t^0) - d\nu_t(\Delta_t^I - \Delta_t^{\hat\theta}) + dF_t + cdt \end{aligned}$$

The second line follows from Lemma 3, the inequality in the third line follows from Condition (i) in Lemma 8, where we take expectations conditional on the event that $dN_t^\theta = dN_t^{\hat\theta} = 0$. The last equality in (4.A.6) holds since

$$q_t^*(\Delta_t^\theta - \Delta_t^{\hat\theta}) = q_t^*(U_t + \Delta_t^\theta - (U_t + \Delta_t^{\hat\theta})) = q_t^*(U_t^j - U_t^i) = \lambda(i-\alpha)(U_t^1 - U_t^0).$$

Punishment is without cost for the principal, and therefore, it is optimal to impose the most severe punishment after an inspection reveals a dishonest report. The severity of punishments is restricted by the limits of enforcement that require the agent's continuation value not to fall below the lower bound $-B < 0$. Therefore, we have

$$\Delta_t^I - \Delta_t^{\hat\theta} = \underbrace{U_t + \Delta_t^I}_{=-B} - (\underbrace{U_t + \Delta_t^{\hat\theta}}_{=U_t^i}) = -(B + U_t^i).$$

Substituting this last equation into Equation (4.A.6) yields

$$dU_t^i = rU_t^i + \lambda(i-\alpha)(U_t^1 - U_t^0) + d\nu_t(B + U_t^i) + dF_t + cdt,$$

which is equal to Condition (H) in Lemma 4. Conversely, if (i) does not hold at some t, then using the same steps as above, the inequality is reversed, so that (H) is violated.

(3.) Substituting Equation (4.A.5) into the obedience constraint (ii) we obtain for each θ_{t-}:

$$(\Delta_t^\theta + \Delta_t^{\hat\theta})(1 - 2\theta_{t-})\alpha\lambda = \alpha\lambda(U_t^1 - U_t^0) \geq c.$$

The last inequality is identical to (O) in Lemma 4. Conversely, if (ii) is violated at some t, then the inequality is reversed, so that (O) is violated. □

Proof of Theorem 1

We solve the full problem in four parts. First, we solve a deterministic impulse control problem setting $\theta = 1$. Second, we show that fines between inspections cannot lower the costs of monitoring. Third, we show that by allowing random transitions in quality, the optimal mechanism is the same as the solution to the deterministic problem whenever $\theta_t = 1$, and for $\theta_t = 0$ it sets a flat fine without inspections. Finally, we verify that the there is no mechanism in non-Markov strategies that performs better.

(1.) *Deterministic impulse control.* We begin by considering the auxiliary control problem where $\theta_t = 1$ (no transitions) and $dF_t = d\mu_t = 0$ whenever $dN_t^I = 0$ (no fines or threats without inspection). This is a standard deterministic impulse control problem with state constrains. Note here that while we rule out transitions in quality, we continue to constrain the evolution of utility to satisfy the incentive-compatibility conditions in Lemma 4. In other words, the agent must be treated as if transitions in quality were possible even though quality is constant over time. Davis (1993) shows that such impulse control problems can be solved iteratively, by solving the sequence of impulse control problems in which the number of inspections is (at most) $n \geq 0$. The value function of this restricted problem will converge to the value function of the original problem as $n \to \infty$.

Fix $t_0 \geq 0$, and for each $\theta \in \{0, 1\}$, denote by $T^\theta(u_t^0, u^1)$ the length of the time interval between t and the time at which U_t^θ hits a boundary, where (u_t^0, u^1) is the profile of promised utilities at t_0 (assuming, as we will later verify, that U_s^1 never hits the lower boundary $-B$.) Since $dF_t = 0$ and $d\mu_t = 0$, the promise-keeping and truth-telling constraints yield a pair of coupled first-order differential equation that have the solution

$$U_t^1 = e^{rt}(u^1 + (1-\alpha)(e^{\lambda t} - 1)(u^1 - u^0)) - c(1 - e^{rt})/r, \tag{4.A.7}$$
$$U_t^0 = e^{rt}(u^0 - \alpha(e^{\lambda t} - 1)(u^1 - u^0)) - c(1 - e^{rt})/r. \tag{4.A.8}$$

Inspection of the truth-telling constraint reveals that U_t^0 is strictly decreasing in t and from the previous two equations, it is easy to see that U_s^0 is strictly increasing in u^0. It thus follows that T^0 is increasing in u^0. In a similar vain, U_s^1 is u-shaped in t and decreasing in u^0. Therefore, T^1 must also be increasing in u^0. Hence, the boundary hitting time T^0 increases in U_0^0 and so for given U_0^1, the principal chooses U_0^0 as large as possible. This means that the obedience constraint $U_0^1 - U_0^1 = c/\lambda\alpha$ must bind. The maximisation problem is thus reduced to finding the optimal initial utility $u := U_t^1$. Define

$$T(u) = \min\{T^0(u - c/\lambda\alpha, u), T^1(u - c/\lambda\alpha, u)\}$$

Step $n = 0$: Suppose the principal cannot perform any interventions. Then the obedience constraint is necessarily violated (all penalties must be enforced independently of true quality) and thus no effort by the agent can be induced. Hence, the value function for the principal is

$$\tilde{V}^0(u) = \bar{V}^0 = \int_t^\infty e^{-(r+\lambda)(s-t)} H \, ds = H/(r+\lambda).$$

Step $n = 1$: Suppose there is only one inspection left to be performed. An increase in u strictly decreases the time U_t^1 hits the upper boundary 0 and strictly increases the time U_t^0 hits the lower boundary $-B$. This is because an upwards shift in the initial promised utility moves up the paths of both U_t^1 and U_t^0. It is thus immediate that to maximise T, the boundary hitting times must be equal, that is:

$$T^0(u - c/\lambda\alpha, u) = T^1(u - c/\lambda\alpha, u).$$

Thus, the optimal initial utility u_1^* is given by the value at which both U^1 and U^0 hit (opposite sides of) the boundary simultaneously. The resulting value function is given by

$$\tilde{V}^1(u) = (1 - e^{-rT(u)})H/r + e^{-rT(u)}(\bar{V}^0 - \kappa).$$

Step $n = 2$: Suppose there are two inspections left to be performed. Let $\phi(s - t, u) := U_s^1$

denote the solution for promised utility given by Equation (4.A.7) with initial condition $(U_t^0, U_t^1) = (u - c/\lambda\alpha, u)$, measuring time relative to the initial instant t. The value function is then

$$\tilde{V}^2(u) = (1 - e^{-rT(u)})H/r + e^{-rT(u)}(\tilde{V}^1(\phi(T(u), u) - \kappa).$$

The optimal initial promised utility u must satisfy the first-order necessary condition

$$\partial_u \tilde{V}^2(u) = rT'(u)e^{-rT(u)}\left[H/r - (\tilde{V}^1(\phi(T(u), u)) - \kappa)\right] + e^{-rT(u)}\partial_u \tilde{V}_1^1(\phi(T(u), u)) = 0.$$

Note here that H/r is the upper bound of the principal's payoff, and thus

$$H/r - (\tilde{V}^1(\phi(T(u), u)) - \kappa) \geq 0. \tag{4.A.9}$$

Thus, $\partial_u \tilde{V}^2(u) = 0$ only if either $T'(u) > 0$ and $\partial_u \tilde{V}^1(\phi(T(u), u)) < 0$ or if $T'(u) < 0$ and $\partial_u \tilde{V}^1(\phi(T(u), u)) > 0$. Suppose the latter is the case. Note that $T'(u) < 0$ means that only the upper boundary is reached, which implies that $\phi(T(u)) = 0$. This in turn implies that $\phi(T(u)) > u_1^*$. But for $u > u_1^*$, we have

$$\partial_u \tilde{V}^1(u) = rT'(u)e^{-rT(u)}\left[H/r - (\tilde{V}^0 - \kappa)\right] < 0,$$

so it cannot be the case that $T'(u) < 0$ and $\partial_u \tilde{V}^1(\phi(T(u), u)) > 0$. Thus, we have $T'(u_2^*) > 0$ and $\partial_u \tilde{V}_1^1(\phi(T(u_2^*), u_2^*)) < 0$. Note also that $T'(u_2^*) > 0$ means that only the lower boundary is reached (higher u increases the hitting time) which in turn implies that $u_2^* < u_1^*$.

Step $n = k$: Suppose there are k interventions left. Now, suppose that $\partial_u \tilde{V}^{k-1}(0) < 0$, as was shown for $k = 2$ in step 2. For $k > 2$, the first-order necessary condition on u is

$$\partial_u \tilde{V}^k(u) = rT'(u)e^{-rT(u)}\left(H/r - \tilde{V}^{k-1}(\phi(T(u), u)) + \kappa\right) + e^{-rT(u)}\partial_u \tilde{V}^{k-1}(\phi(T(u), u)) = 0.$$

If $T'(u_k^*) < 0$, then only the upper boundary is reached, which implies that $\phi(T(u_k^*)) = 0$ and $\partial_u \tilde{V}^{k-1}(0) < 0$, a contradiction. Thus, the solution must have $T'(u_k^*) > 0$ and $\partial_u \tilde{V}^{k-1}(0) < 0$. Moreover, we have $\partial_u \tilde{V}^k(0) < 0$ since $T'(0) < 0$ and $\partial_u \tilde{V}^{k-1}(0) < 0$, which proves the induction hypothesis for $k + 1$.

There is a value $\underline{u}$ such that $u \leq \phi(T(u), u)$ for all $u \geq \underline{u}$ (this follows from the laws of evolution of U^0, U^1). We show that $u_k^* \geq \underline{u}$ for all k. For $k = 1$, we know $u_2^* < u_1^*$, so that $u_2^* \geq \underline{u}$. Let k be the first index at which $u_k^* < \underline{u} < u_{k-1}^*$. Since $u_k^* < \underline{u}$, we have $u_k^* > u_{k-1}^*$, but by hypothesis, $u_{k-1}^* > \underline{u}$, a contradiction. We thus have $u_k^* \geq \underline{u}$, which implies that $\underline{u} \leq u_k^* < u_{k-1}^*$ for all $k > 1$. Therefore, the sequence $(u_k^*)_k$ is bounded and strictly decreasing, and thus by the monotone convergence theorem, it has a limit u^* with $u^* > \underline{u}$.

(2.) *No fines between inspections.* Consider now a mechanism $\{F_t, N_t^I\}_t$ in which $\{F_t\}_t$ and $\{\mu_t\}_t$ are any measurable processes with $dF_t, d\mu_t \geq 0$. Since U_t^0 is decreasing over time, it is clear that $d\mu_t = 0$ for all t since otherwise, the time until U_t^0 reaches the enforcement boundary decreases, thus increasing monitoring costs. Starting from the original mechanism with arbitrary fines, we show that there exists a mechanism without fines between inspections that generates the same costs to the principal.

In the original mechanism with arbitrary fines, suppose $t < T$ are two consecutive inspection dates. Let $u = (u^0, u^1)$ be the utility of the agent at time t, and denote by $V(U)$ the principal's expected monitoring costs at state $U = (U^0, U^1)$. Since the principal incurs the cost only at the time of inspections, the monitoring costs for the principal at time T can be written as

$$V(u) = e^{-rT}V(U_T)$$

where $U_T^\theta = u^\theta + \int_t^T dU_s^\theta$ for each $\theta \in \{0, 1\}$.

Consider now an alternative mechanism $\{\hat{F}_t, \hat{N}_t^I\}_t$ that is identical to the original mechanism at any time $s < t$ and $s \geq T$. However, we set $d\hat{F}_s = 0$ at all $s \in (t, T)$ and $d\hat{F}_t = dF_t + F_T - F_t$ at time t. Denote by $\hat{U}_s^\theta$ the promised utility to the agent in this alternative mechanism, and let the expected monitoring costs for the principal in the new mechanism be denoted by $\hat{V}(U)$. The monitoring costs generated in these mechanisms are equal, since

$$\hat{V}(u + d\hat{F}_t) = e^{-rT} V(\hat{U}_T) = e^{-r} V\left(u + F_T - F_t + \int_t^T d\hat{U}_s\right)$$
$$= e^{-rt} V\left(u + \int_t^T (d\hat{U}_s + dF_s)\right) = e^{-rT} V(U_T) = V(u).$$

Since this modification can be performed for any pair of consecutive inspection dates, the mechanism constructed in part (1.) of this proof is optimal among mechanisms with arbitrary fine processes $\{dF_t\}_{t\geq 0}$.

(3.) *Quality transitions.* We now consider the original model, in which quality follows a two-state Markov chain. In this case, the process is piecewise deterministic because payoffs evolve deterministically between quality transitions. Suppose $u = (u^0, u^1)$ is the optimal initial utility pair in state $\theta = 1$. We show that in any contract, it is possible to have $U_t = \bar{U}$ at all t when $\theta = 0$ without performing inspections.

The proof is constructive. Suppose $\bar{U}$ is the initial utility when $\theta = 1$. Since $dU_t^0 \leq 0$ by the truth-telling conditions, we have $U_t^0 \leq \bar{U}$. Set $dF_t = \bar{U}^0 - U_t^0$ and $d\mu_t = \bar{U}^0 - U_t^0 - (\bar{U}^1 - U_t^1)$. By promise keeping and truth telling, we have $dU_t^1 - dU_t^0 \geq 0$, and hence $d\mu_t \geq 0$. We thus have $U_{t+}^i = \bar{U}^i$ for $i = 0, 1$. Therefore, if the initial utility in state $\theta = 1$ is $\bar{U}$, it can be reset to $\bar{U}$ when the state switches to $\theta = 0$. Next, while $\theta = 0$, set $dF_t = -r\bar{U}^0 + \alpha\lambda(\bar{U}^1 - \bar{U}^0)$ and $d\mu_t = c + (r + \lambda)(\bar{U}^1 - \bar{U}^0)$. Substituting into the promise-keeping and truth-telling constraints, it follows that $dU_t^i = 0$ for each $i = 0, 1$, and thus we can construct a contract with $U_t = \bar{U}$ for all t at which $\theta = 0$.

(4.) *General mechanisms.* Parts (1.)-(3.) demonstrate that the mechanism described in the theorem is an optimal Markovian mechanism. It remains to show that no (non-Markovian) mechanism can do better. Let $V^{\theta_t}(U)$ denote the expected costs for the principal in our mechanism that delivers the agent with promised payoffs of $U = (U^0, U^1)$. We show that the expected costs in state θ_t from any incentive-compatible mechanism that delivers the initial expected value $U_0 = (U_0^0, U_0^1)$ to the agent cannot lie below $V^{\theta_t}(U_0)$. Since the inspection cost κ does not depend on the state value prior to inspection, we can apply Theorem 54.28 from Davis (1993, p. 242) to conclude that the value function V the unique continuous and bounded function that solves the quasi-variational inequality

$$\mathcal{U}V^\theta(u) - rV^\theta(u) \geq 0,$$
$$WV^\theta(u) - V^\theta(u) \geq 0, \text{ and}$$
$$\left(\mathcal{U}V^\theta(u) - rV^\theta(u)\right)\left(WV^\theta(u) - V^\theta(u)\right) = 0,$$

on the state space $\left\{(\theta, u^0, u^1) \colon \theta \in \{0, 1\}, (u^0, u^1) \in [-B, 0]^2, u^1 - u^0 \geq \frac{c}{\lambda\alpha}\right\}$. Here, $\mathcal{U}$ is the extended generator of the piecewise deterministic Markov process which is defined by the relationship[9]

$$\mathbb{E}_{\theta,u} V^{\theta_t}(U_t) = V^{\theta}(u) + \mathbb{E}^{\theta,u} \int_0^t \mathcal{U} V^{\theta_s}(u_s) ds$$

in case no inspection occurs before t, and W is the expected total cost at the time of an inspection.

$$W V^{\theta} = \min_{u_0,u_1} V^{\theta}(u_0, u_1) + \kappa.$$

Consider an arbitrary incentive-compatible mechanism with inspection process $\{dN_t^I\}_t$ and define the total expected cost at time t by

$$G_t = \int_0^t e^{-rs} \kappa dN_s^I + e^{-rt} V^{\theta_t}(U_t).$$

For $t = 0$, we have $G_0 = V^{\theta_0}(U_0)$. For $t > 0$, we can represent G_t by the differential formula (see Theorem 31.3 in Davis, 1993, p. 83) as

$$\mathbb{E}_s[G_t] - G_s = \int_s^t e^{-r(z-s)} \left(\mathcal{U} V^{\theta_z}(U_z) - r V^{\theta_z}(U_z) \right) dz + \mathbb{E}_s \left[\int_s^t e^{-r(z-s)} \left(W V^{\theta_z}(U_z) - V^{\theta_z}(U_z) \right) dN_z^I \right].$$

By the variation inequality above, both integrals are non-negative so that the process $(G_t)_t \geq 0$ is a submartingale and $E_0 [G_t] \geq G_0$ for any $t \geq 0$. In particular, taking the limit as t approaches infinity, we get $E_0 \left[\int_0^\infty e^{-rs} \kappa dN_s^I \right] = E_0 [\lim_{t \to \infty} G_t] \geq G_0 = V^{\theta_0}(U_0)$. Hence, any incentive-compatible maximal-compliance mechanism leads to weakly higher expected inspection costs.

□

Proof of Lemma 5

Define

$$\Psi(T) \equiv (B - c/r)\left(1 - e^{-rT}\right) - c/(\lambda\alpha) e^{\lambda T}\left(e^{rT} - \alpha\right) + c/(\lambda\alpha)(1 - \alpha), \quad (4.A.10)$$

so that $T^* = \inf\{T > 0 : \Psi(T) = 0\}$. This exists and is unique whenever our feasibility assumption $B > c\frac{r+\lambda}{r\lambda\alpha}$ is satisfied (Ψ is increasing from 0 at $T = 0$ and crosses 0 from above exactly once). The function Ψ is continuously differentiable in all parameters and in T on a neighbourhood of T^*. By the implicit function theorem we have

$$\frac{\partial T^*}{\partial x} = - \left. \frac{\Psi_x}{\Psi_T} \right|_{T=T^*},$$

for all parameters $x \in \{B, c, \alpha, \lambda\}$, where Ψ_x denotes the partial derivative of Ψ with respect to x. As mentioned above, $\Psi(T)$ crosses 0 from above at $T = T^*$ so that $\Psi_T|_{T=T^*} < 0$. Hence, for all parameters, we have

$$sign\left(\frac{\partial T^*}{\partial x}\right) = sign\left(\Psi_x|_{T=T^*}\right).$$

The first two items of Lemma 5 follow immediately as Ψ is increasing in B and decreasing in c everywhere.

Likewise for the third item, note that

$$\Psi_\alpha = \frac{c}{\lambda\alpha^2}\left(e^{(r+\lambda)T} - 1\right) > 0,$$

[9] See Davis, 1993, pp. 27-33.

so that T^* is increasing in α.

For the fourth item, describing the change of T^* in λ, consider Ψ in (4.A.10) as $\lambda \searrow \frac{cr}{Br\alpha - c}$, which is the lower bound on λ such that the feasibility assumption $B > \frac{c(r+\lambda)}{r\lambda\alpha}$ is fulfilled. $\Psi = 0$ is then equivalent to

$$(B - \frac{c}{r})(1 - e^{-rT}) - \left(B - \frac{c}{r\alpha}\right)\left(e^{\lambda T}(e^{rT} - \alpha) - (1-\alpha)\right) = 0.$$

This can only be fulfilled at $T = 0$, as we have $B > \frac{c}{r\alpha} > \frac{c}{r}$ and for all $T > 0$,

$$0 > -\left(e^{\lambda T}(e^{rT} - \alpha) - (1-\alpha)\right) > \frac{1}{\alpha}\left(e^{\lambda T}(e^{rT} - \alpha) - (1-\alpha)\right).$$

Hence, T^* is initially increasing in λ. Finally, consider Ψ in (4.A.10) to see that $T^*(\lambda) \overset{\lambda\to\infty}{\longrightarrow} 0$. In particular,

$$\lim_{\lambda\to\infty} \frac{e^{(r+\lambda)T^*(\lambda)}}{\lambda} = 0.$$

This implies that $\lambda T^*(\lambda)$ is either finite or grows at lower than logarithmic rate as λ becomes arbitrarily large. Thus, $T^*(\lambda)$ must go to 0. $\square$

Proof of Lemma 6

The first two items follow immediately from the previous lemma as C^H is decreasing in T.

Consider the total derivative of costs C^H w.r.t. α:

$$\begin{aligned}\frac{d}{d\alpha}C^H = \Big[&(2\alpha-1)\lambda^2 + e^{(r+\lambda(1-\alpha))T^*}\left((T^*\lambda(r+\alpha\lambda)(r+\lambda(1-\alpha)) - (2\alpha-1)\lambda^2\right) \\ &- T^*_a \cdot (r+\lambda(1-\alpha))^2(r+\alpha\lambda)\Big]\frac{1}{(e^{(r+\lambda(1-\alpha))T^*}-1)^2 r(r+\lambda)}.\end{aligned} \tag{4.A.11}$$

The change in inspection costs caused by varying α contains the negative effect through the increase in T^* and an effect on the environment contained in the first terms of the squared bracket. The second effect captures the change in relative probability of high reports as well as the variability of the state, both of which determine how often the deadline is reached without previously changing to state L. To see that it is always positive, verify that it is 0 at $T = 0$ and increasing in T.

We establish the second item of the lemma: there exists $\underline{\alpha} > \frac{c(r+\lambda)}{Br\lambda}$ such that C^H is decreasing in α for all $\alpha < \underline{\alpha}$. Note that as $\alpha \searrow \frac{c(r+\lambda)}{Br\lambda}$, $T^*(\alpha) \searrow 0$. The squared bracket in (4.A.11) converges to $-T^*_\alpha|_{\alpha\searrow\frac{c(r+\lambda)}{Br\lambda}} \cdot r(r+\lambda)^2$. T^*_α is strictly positive for $\alpha > \frac{c(r+\lambda)}{Br\lambda}$. By continuity, the derivative must be negative for all α smaller than $\underline{\alpha} > \frac{c(r+\lambda)}{Br\lambda}$.

Last, we show that there exists $\bar{\alpha}$ such that C^H is increasing in α for all $\alpha > \bar{\alpha}$ and that $\bar{\alpha} < 1$ whenever $\frac{Br-c}{cr}$ is large enough. Consider the squared bracket in (4.A.11) as $\alpha \nearrow 1$. This is equal to

$$\lim_{\alpha\to1} -\lambda^2\left[e^{rT(\alpha)} - 1\right] + \left(\lambda T(\alpha) - rT'(\alpha)\right)e^{rT(\alpha)}.$$

From $\Psi|_{\alpha=1} = 0$ we get $e^{(r+\lambda)T^*(\alpha)} \overset{\alpha\to 1}{\longrightarrow} \frac{Br-c}{cr}\lambda$. Further,

$$
\begin{aligned}
T^*_\alpha|_{\alpha\to 1} &= \lim_{\alpha\nearrow 1}\left(-\left.\frac{\lambda(Br-c)(1-e^{-rT})+cr(e^{\lambda T}-1)}{\lambda\alpha(Br-c)re^{-rT}-(r+\lambda)cre^{(r+\lambda)T}+\lambda\alpha cre^{\lambda T}}\right|_{T=T(\alpha)}\right)\\
&= \frac{\lambda\frac{(Br-c)}{cr}\left(\left(\frac{Br-c}{cr}\lambda\right)^{\frac{r}{r+\lambda}}-1\right)+\left(\left(\frac{Br-c}{cr}\lambda\right)-\left(\frac{Br-c}{cr}\lambda\right)^{\frac{r}{r+\lambda}}\right)}{-\lambda\frac{(Br-c)}{cr}r+(r+\lambda)\left(\frac{Br-c}{cr}\lambda\right)^{\frac{2r+\lambda}{r+\lambda}}-\lambda\left(\frac{Br-c}{cr}\lambda\right)}\\
&= \frac{\left(\frac{Br-c}{cr}\lambda\right)^{\frac{2r+\lambda}{r+\lambda}}-\left(\frac{Br-c}{cr}\lambda\right)^{\frac{r}{r+\lambda}}}{(r+\lambda)\left(\left(\frac{Br-c}{cr}\lambda\right)^{\frac{2r+\lambda}{r+\lambda}}-\frac{(Br-c)}{cr}\lambda\right)}.
\end{aligned}
$$

Inserting into (4.A.11) at $\alpha = 1$ and defining $\chi = \frac{(Br-c)}{cr}\lambda > 1$, we see that the deterministic inspection costs are increasing in α if and only if

$$\lambda^2 + \chi\left(-\lambda^2 + \frac{\lambda}{r+\lambda}\ln(\chi) - \frac{\chi^{\frac{2r+\lambda}{r+\lambda}} - \chi^{\frac{r}{r+\lambda}}}{\chi^{\frac{2r+\lambda}{r+\lambda}} - \chi}\right) > 0.$$

As χ grows large (for example as B increases), the fraction in the bracket approaches 1, so the second term grows arbitrarily large. Therefore, we have that for χ large enough, there exists $\bar{\alpha} < 1$ such that the deterministic costs are increasing in α for all $\alpha > \bar{\alpha}$.

In the case of λ, the first result, that the costs decrease initially in λ, is shown analogously to the corresponding result in the case of α. To see that the costs become arbitrarily large in the limit, recall from the previous proof that $\lambda T^*(\lambda)$ grows to ∞ at lower than logarithmic rate. The total costs in the limit are thus given by

$$\lim_{\lambda\to\infty} C^H = \frac{(1-\alpha)\alpha}{r}\lim_{\lambda\to\infty}\frac{\lambda}{e^{(1-\alpha)\lambda T^*(\lambda)}} = \infty.$$

□

References

Antinolfi, G. and Carli, F. (2015). Costly Monitoring, Dynamic Incentives, and Default, *Journal of Economic Theory* **159**: 105 – 119.

Avenhaus, R., Von Stengel, B. and Zamir, S. (2002). Inspection games, *Handbook of Game Theory with Economic Applications* **3**: 1947–1987.

BaFin (2016). Richtlinie zur Durchführung und Qualitätssicherung der laufenden Überwachung der Kredit- und Finanzdienstleistungsinstitute durch die Deutsche Bundesbank.
URL: *https://www.bafin.de/dok/7852628*

Bassetto, M. and Phelan, C. (2008). Tax Riots, *The Review of Economic Studies* **75**(3): 649–669.

Becker, G. S. (1968). Crime and Punishment: An Economic Approach, *Journal of Political Economy* **76**(2): 169–217.

Ben-Porath, E. and Kahneman, M. (2003). Communication in Repeated Games with Costly Monitoring, *Games and Economic Behavior* **44**(2): 227 – 250.

Board, S. and Meyer-ter-Vehn, M. (2013). Reputation for quality, *Econometrica* **81**(6): 2381–2462.

Bond, P. and Hagerty, K. (2010). Preventing Crime Waves, *American Economic Journal: Microeconomics* **2**(3): 138–159.

Border, K. C. and Sobel, J. (1987). Samurai Accountant: A Theory of Auditing and Plunder, *The Review of Economic Studies* **54**(4): 525–540.

Carey, P. and Guest, R. (2000). Determining the Optimal External Audit Interval for Private (and Family-Controlled) Companies, *Journal of Accounting, Auditing & Finance* **15**(4): 439–458.

Chang, C. (1990). The Dynamic Structure of Optimal Debt Contracts, *Journal of Economic theory* **52**(1): 68–86.

Chang, C. F., Steinbart, P. J. and Tuckman, H. P. (1993). Are Medicare audits random or predictable?, *Journal of Accounting and Public Policy* **12**(2): 135 – 154.

Davis, M. H. (1993). *Markov Models and Optimization*, Vol. 49 of *Monographs on Statistics and Applied Probability*, Chapman & Hall.

Dye, R. A. (1986). Optimal Monitoring Policies in Agencies, *The Rand Journal of Economics* pp. 339–350.

EU (2016). Directive (EU) 2016/680 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data by competent authorities for the purposes of the prevention, investigation, detection or prosecution of criminal offences or the execution of criminal penalties, and on the free movement of such data, and repealing Council Framework Decision 2008/977/JHA. 6. L 119.

European Commission (2019). Commission Recommendation (EU) 2019/1318 on internal compliance programmes for dual-use trade controls under Council Regulation (EC) No 428/2009. European Commission, 30 July 2019.

Fernandes, A. and Phelan, C. (2000). A Recursive Formulation for Repeated Agency with History Dependence, *Journal of Economic Theory* **91**(2): 223–247.

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

Hughes, J. (1977). Optimal Internal Audit Timing, *The Accounting Review* **52**(1): 56–68.

Kamada, Y. and Rao, N. (2018). Strategies in Stochastic Continuous-Time Games. Working Paper, Haas School of Business, University of California, Berkeley.

Kaplan, R. S. (1969). Optimal Investigation Strategies with Imperfect Information, *Journal of Accounting Research* **7**(1): 32–43.

Kim, S.-H. (2015). Time to Come Clean? Disclosure and Inspection Policies for Green Production, *Operations Research* **63**(1): 1–20.

Krasa, S. and Villamil, A. P. (2000). Optimal Contracts When Enforcement is a Decision Variable, *Econometrica* **68**(1): 119–134.

Last, G. and Brandt, A. (1995). *Marked Point Processes on the Real Line: The Dynamical Approach*, Springer.

Majone, G. (1994). The Rise of the Regulatory State in Europe, *West European Politics* **17**: 77–101.

Monnet, C. and Quintin, E. (2005). Optimal Contracts in a Dynamic Costly State Verification Model, *Economic Theory* **26**(4): 867–885.

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

Osaki, S. (2002). *Stochastic Models in Reliability and Maintenance*, Springer.

Popov, L. (2016). Stochastic Costly State Verification and Dynamic Contracts, *Journal of Economic Dynamics and Control* **64**: 1–22.

Ravikumar, B. and Zhang, Y. (2012). Optimal Auditing and Insurance in a Dynamic Model of Tax Compliance, *Theoretical Economics* **7**(2): 241–282.

Reinganum, J. F. and Wilde, L. L. (1985). Income Tax Compliance in a Principal-Agent Framework, *Journal of Public Economics* **26**(1): 1–18.

The Economist (2020). The Brussels effect. The EU wants to set the rules for the world of technology, *The Economist,* **Feb 20, 2020**.

Townsend, R. M. (1979). Optimal Contracts and Competitive Markets with Costly State Verification, *Journal of Economic Theory* **21**(2): 265–293.

Varas, F., Marinovic, I. and Skrzypacz, A. (2020). Random Inspections and Periodic Reviews: Optimal Dynamic Monitoring, *Review of Economic Studies* (forthcoming).

Wang, C. (2005). Dynamic Costly State Verification, *Economic Theory* **25**(4): 887–916.

Webb, D. C. (1992). Two-Period Financial Contracts with Private Information and Costly State Verification, *The Quarterly Journal of Economics* **107**(3): 1113–1123.

Zhang, Y. (2009). Dynamic Contracting with Persistent Shocks, *Journal of Economic Theory* **144**(2): 635–675.