Essays in Learning and Information Design

Carlo Antonio Cabrera*

*A thesis submitted to the Department of Economics of the London School of Economics and Political Science for the degree of Doctor of Philosophy, London, October 2018

To Ely

Declaration

I certify that the thesis I have presented for examination for the degree of Doctor of Philosophy to the London School of Economics and Political Science is solely my own work.

I warrant that this authorisation does not, to the best of my belief, infringe the rights of any third party.

I declare that my thesis consists of 24,999 words.

Statement of inclusion of previous work

I can confirm that chapter 3 is a result of previous study for an MRes in Economics which I undertook at the London School of Economics and Political Science and completed in 2014.

Acknowledgements

I owe my gratitude to a number of people.

To my parents, who have given me nothing but endless support.

To Ely, who had the patience to wait for me to finish writing this thesis. I'm sorry it took so long.

To Alkis, who helped me realise that research doesn't have to be lonely.

To Thomas, who always had creative approaches to any problem.

To Gerard Padro, who gave me helpful supervision and guidance in my second year.

To Erik, without whom I probably would not even be at the LSE.

To the rest of the theory faculty: Ronny, Gilat, Francesco, Matt, and Andrew, for always being so generous with your time.

To my supervisor, Balázs, for his help, encouragement, and for being the most insightful person I know.

Thank you all so very much. I couldn't have done this without you.

Abstract

We re-visit three classical models in information economics.

The first chapter studies the screening problem for a seller who owns a single good, and a buyer whose valuation for the good is their private information. We allow for the seller to acquire information at some cost about the buyer's value, in addition to her choice over the probability of trade and the transfer. The seller thus chooses a Blackwell experiment for each announcement that the buyer makes in a direct revelation mechanism. More informative experiments are more costly. Under mild conditions, there are always optimal mechanisms where the seller acquires coarse information about the buyer. In particular, it is always optimal for the seller to choose an experiment that consists of no more than four signals. When the buyer has only two possible values, the same holds for experiments that consist of at most three signals.

The second chapter examines information disclosure in a setting of strategic experimentation. A group of agents continuously and independently choose between a safe arm and risky arms of the same type. When the arms reveal good news, we are able to achieve efficiency in a class of simple information disclosure mechanisms when the agents are initially optimistic enough about their risky arms, but only when there are not too many agents. When the reverse is true, the mechanism must be transparent. Thus, there is a tradeoff between transparency and efficiency. This tradeoff does not exist in the case of bad news.

In the final chapter, we borrow insights from social learning theory to understand why institutions have persistent effects. We adapt the classical model in a minimal fashion to accommodate a role for institutions, and demonstrate that social learning is one plausible mechanism of persistence.

Contents

1 Screening with Information Acquisition

1.1 Introduction

Modern technology has brought unprecedented ease to the access of information. This is as true for firms as it is for consumers. Firms can use cookies to track the Internet activity of current and potential customers. They can access GPS data in smartphones to track consumer movement.[1] They can purchase information from data brokers, who collect data about consumers from a wide variety of sources.[2] The ability of firms to gather information about consumers has grown to such an alarming extent as to incite aggressive regulatory response. For example, in early 2018, the European Union enacted the General Data Protection Regulation (GDPR), which places significant restrictions on the acquisition, use, and handling of information on consumers in Europe by firms.

Acquiring information is costly. Firms must first gather data, whether on their own or through a data broker, and invest resources into transforming this data into useful information. Firms also need to do this in a manner that respects regulation like the GDPR. On the other hand, useful information is incredibly valuable. Otherwise, we would not have observed such explosive growth in the data brokerage industry.[3] This is also why some of the largest corporations in the world, such as Google and Facebook, are in the business of information. One is then led to ask: how much information is optimal for firms to acquire? This question is of natural economic interest. Moreover, our understanding of this question can inform how we regulate information acquisition activity. We aim to address this question here.

Consider the classic model of screening with a seller who possesses a single indivisible good and a buyer whose value is their private information.[4] Suppose that, aside from being able to design the selling mechanism, our seller could also acquire information about the buyer at some cost. More precisely, in a direct revelation mechanism, the seller

[1] For example, Google Maps Timeline keeps detailed records of a smartphone's location history.

[2] A summary of the activities of data brokers can be found in Bergemann and Bonatti (2015). According to US Senator John D. Rockefeller, in 2012, the data brokerage industry generated revenue of \$156 billion, 'a sum more than twice the size of the entire intelligence budget of the United States Government'. See Rieke et al. (2016).

[3] See footnote 2.

[4] We assume that the buyer maximises quasilinear expected utility, and that the buyer's value lies in some finite subset of $\mathbb{R}_+$.

chooses a (finite) Blackwell experiment (1951) as well as the probability of trade and transfer for each type announcement of the buyer. The probability of trade and transfer is thus not only a function of the buyer's type announcement, but it now depends on the signal generated by the seller's chosen experiment as well. The seller's payoff is her revenue less the cost of acquiring information. We assume that the seller's cost is non-decreasing in the Blackwell-dominance order. Additionally, to rule out the full extraction of surplus at arbitrarily small costs, we assume that both the buyer and the seller face limited liability constraints: the buyer cannot pay arbitrarily large amounts, nor the seller pay arbitrarily large transfers to the buyer. We also impose incentive and participation constraints. What are the properties of the optimal mechanism?

Our first main result states that there always exists[5] an optimal mechanism with a simple structure. More specifically, there is always an optimal mechanism where the seller never acquires more than four signals. Furthermore, in this simple mechanism, the seller never chooses an interior probability of trade or an interior transfer after each signal. Hence, there always exists an optimal mechanism where the seller acquires coarse information about the buyer. This result follows from two simple observations. Take any optimal mechanism, and notice that one can always transform this into a mechanism where the probabilities of trade and transfers are extremal, as described above, by modifying the signal probabilities in a way that achieves the same revenue, and respects the incentive and participation constraints. This transformation is a garbling of the seller's original information structure, and therefore cannot increase the seller's cost. Thus, for each experiment, there are only four relevant outcomes. The seller can then garble each experiment further to reduce the number of signals to four.

Our second main result is specialised to the case where the buyer has only two possible types. In this setting, it is always optimal for the seller to never acquire information when the buyer declares that they are the high type. When the buyer declares that they are the low type, then it is optimal for the seller to acquire no more than three signals, and choose extremal probabilities of trade and transfers. That seller acquires no information about the high type follows from the fact that the low type never mimics the high type, and that acquiring information is useful only for relaxing incentive constraints. On the other hand, when designing the mechanism for the low type, all the seller cares about is the likelihood ratio of each signal. In particular, the solution to her revenue maximisation problem for a fixed signal structure can be characterised by a pair of cutoffs on the likelihood ratios. One cutoff determines whether the seller should allocate the good, while the other determines whether to charge the largest transfer or the smallest.[6] Therefore, the two cutoffs partition the space of likelihood ratios into at most three regions, which guarantees that we can always garble the information structure into one where the seller acquires no more than three signals.

[5] We have to impose further mild conditions to guarantee the existence of an optimum. See proposition 1.1.

[6] There is a minor complication in this argument, because it leaves unspecified what the seller should do if the likelihood ratio is exactly at a cutoff. This case can be dealt with in exactly the same way as in the argument given for the first main result.

These two results are appealing, because they reveal that the problem has an astonishingly simple structure, despite the generality of the assumed cost function. Moreover, these results suggest that the regulations that limit information acquisition need not entail any efficiency losses beyond the cost of implementation, contrary to classical predictions of economic theory. In particular, it is not necessarily the case that there is a tradeoff between (neoclassical) efficiency and privacy.

Our third result explores the relaxation of the limited liability constraint. Under mild assumptions on the cost function, without limited liability, the seller can always extract full surplus at vanishingly small cost. Hence, any optimal mechanism, if one exists, must extract surplus fully at no cost to the seller. This result continues to hold even if only the buyer faces limited liability, but the seller is allowed to make arbitrarily large payments to the buyer. This result underscores the need for limitations on the seller if one is concerned about issues regarding the distribution of surplus between consumers and firms: the ability to acquire information can strongly skew this distribution in the seller's favour.

Finally, we address the tightness of the bounds on the number of signals stated above. Is it possible that there is always an optimal mechanism simpler than the ones described above? In general, the answer is no, at least in the binary-type case. This is established by constructing an example where the seller cannot do better by acquiring less than three signals.

The rest of the paper proceeds as follows. Section 1.2 discusses the related literature, and section 1.3 defines the model. 1.4 covers the main results, while section 1.5 is a discussion of the limited-liability assumption and the tightness of the main results. Section 1.6 concludes.

1.2 Related Literature

This paper extends the classic screening model, which has become a workhorse in economics, used to analyse price discrimination, credit rationing, optimal income taxation, labour contracts, and optimal regulation. An overview of the model and its applications can be found in Bolton and Dewatripont (2005). In the literature on price discrimination, Bergemann et al. (2015) is especially closely related to what we study here. In particular, their analysis characterises the welfare implications as one varies the mechanism in our setting. One crucial difference is that they treat the information structure as exogenous, whereas we treat it as an object of choice.

There is also a large literature on information in mechanism design, as surveyed in Bergemann and Välimäki (2006). However, the bulk of the literature on adverse selection focuses on information acquisition acquisition by agents, as in, for instance, Bergemann and Välimäki (2002); Gershkov and Szentes (2009); Milgrom (1981); Roesler and

Szentes (2017), or information disclosure by the principal, cf. Bergemann and Pesendorfer (2007); Eső and Szentes (2007); Dworczak (2017). On the other hand, there is a relatively large literature on information acquisition by the principal in settings with moral hazard, where the principal designs a monitoring structure for the agent. See, for example, Georgiadis and Szentes (2018) and the references therein. Our setting is one of adverse selection where the principal optimally acquires information.

There is an older literature that studies settings of adverse selection where the principal acquires information that originates in Townsend (1979), who studies a model of costly state verification. A common assumption in this literature is that the principal has a binary choice of whether to pay some fixed cost in order to verify the state, and the principal learns the state perfectly whenever she decides to pay this cost. We differ from this literature by allowing our principal design a flexible information structure where she can choose to learn partial information about the state. There is also a more recent paper by Strulovici and Siegel (2018), who study the design of judicial mechanisms, and part of that design involves acquiring information about whether a defendant is guilty of a crime. Their design of judicial mechanisms share some of the features we find here. In particular, their state is binary (i.e. a defendant is either guilty or innocent), and they find that it is optimal not to acquire information whenever a defendant pleads guilty. In a similar vein, with two types, we find that it is optimal not to acquire information about the high type. However, they make no assumptions about the cost of acquiring information, and thus have little to say about optimal information structure. They focus instead on characterising the features of the optimal judicial mechanism for a fixed information structure.

Finally, this paper is related to a disparate literature on costly information acquisition studied in variety of different settings. The study of rational inattention is particularly emblematic of this area. See, for instance, Matějka and McKay (2015), who study a discrete choice decision problem where the decision maker can acquire information at a cost proportional to the entropy of the information that she acquires. Other examples include Argenziano et al. (2016), who study a Crawford-Sobel (1982) type cheap talk setting where the receiver can acquire information about the state. Matysková (2018) does a similar exercise when the sender has commitment power, extending the model of Bayesian persuasion in Kamenica and Gentzkow (2011).

1.3 Model

1.3.1 Environment

There is a single buyer (they) with unit demand, and a single seller (she) with one indivisible good. The seller has full commitment. The buyer's value for the good is their private information and takes values in a finite set $V \subset (0,1)$.[7] The buyer's value is thus their (Harsanyi) type. The buyer maximises expected utility that is quasi-linear in transfers. That is, a buyer with value v who pays a transfer t has utility $v - t$ when they receives the good, and has $-t$ otherwise. The seller's beliefs over the buyer's value is given by $p \in \Delta(V)$. Assume that the seller places strictly positive probability on each element of V. Her payoffs will be described below.

1.3.2 Mechanisms

A *mechanism* $(\mathcal{M}, \mathcal{S})$ is a pair consisting of a *trading strategy*

$$\mathcal{M} = \left\{ \left(x^i, t^i \right) \right\}_{i \in V}$$

and an *information structure* $\mathcal{S} = \left\{ \mathcal{S}^i \right\}_{i \in V}$, where

$$\mathcal{S}^i = \left(S^i, \left\{ \sigma^i_j \right\}_{j \in V} \right).$$

An information structure $\mathcal{S}$ is a collection $\left\{ \mathcal{S}^i \right\}_{i \in V}$ of (Blackwell) experiments. An experiment $\mathcal{S}^i$ consists of a finite set of signals S^i, and a collection $\left\{ \sigma^i_j \right\}_{j \in V} \subset \Delta\left(S^i\right)$ of probability distributions over signals in S^i. That is, σ^i_j is a probability distribution on S^i for each i and j in V. We will denote the probability of the signal $s \in S^i$ under the distribution σ^i_j by $\sigma^i_j(s)$. We interpret $\sigma^i_j(s)$ as the probability that signal s is realised when the buyer with true value j announces that their value is i.[8]

A trading strategy $\mathcal{M}$ is a collection of pairs of functions $\left(x^i, t^i\right)$ on S^i with $x^i : S^i \to [0,1]$ and $t^i : S^i \to [0,1]$ for each i in V.[9] As will be made clearer below, for each signal realisation in S^i, x^i determines the probability that the good is sold to the buyer, and t^i determines the transfer the buyer must pay to the seller, independently of whether the object is traded.

1.3.3 Timing and Payoffs

The seller commits to a mechanism $(\mathcal{M}, \mathcal{S})$, which the buyer observes. The seller's choice of mechanism induces a decision problem for the buyer, where the buyer chooses

[7] We will occasionally abuse notation by also referring to V as a set that indexes the set of possible values of the buyer. In particular, we will treat i and j as members of V and, respectively, synonyms of v_i and v_j.

[8] Our definition of a Blackwell experiment differs slightly from the classical definition, which takes the joint distribution between signals and values as primitive. In our formulation, $\sigma^j_i(s)$ is the probability of the signal s *conditional* on value i. This equivalent definition will be easier to work with in our setup.

[9] The fact that $t^i \in [0,1]$ is a substantive restriction. This is equivalent to *both* the seller and buyer facing limited liability. The seller cannot make transfers to the buyer, and the buyer cannot pay transfer greater than 1. A restriction of this type is necessary for our maximisation problem to be well-defined.

to opt in or out of the mechanism. If the buyer opts out, both players receive a payoff of zero. If the buyer opts in, they must announce a value to the seller.[10] A signal $s \in S^i$ is then realised according to the distribution σ^i_j, where i is the buyer's announced value and j is their true value. The good is subsequently traded with probability $x^i(s)$, and the buyer pays the seller a transfer $t^i(s)$.

Hence, a buyer with value v_j who opts in to the mechanism and announces v_i enjoys a payoff of[11]

$$U^i_j(\mathcal{M}, \mathcal{S}) = \sum_{s \in S^i} \sigma^i_j(s) \left[v_j x^i(s) - t^i(s) \right].$$

Whenever the buyer opts in and announces their type truthfully, the seller's revenue is given by

$$R(\mathcal{M}, \mathcal{S}) = \sum_{i \in V} \left[p_i \sum_{s \in S^i} \sigma^i_i(s)\, t^i(s) \right].$$

If the seller sought only to maximise revenue, then the problem is trivial. Hence, the seller must pay a cost for her choice of information structure $\mathcal{S}$. We describe this cost in the next section.

1.3.4 The Cost of Information

An information structure $\mathcal{S}$ is a collection of Blackwell experiments. Let $\mathcal{S}^i = \left(S^i, \left\{ \sigma^i_j \right\}_{j \in V} \right)$ be the i^{th} (Blackwell) experiment in $\mathcal{S}$, and denote the Blackwell-dominance order by $\succeq$. Say that $\mathcal{S}$ *(point-wise) Blackwell-dominates* $\hat{\mathcal{S}}$ if $\mathcal{S}^i \succeq \hat{\mathcal{S}}^i$ for each i in V. In words, $\mathcal{S}$ dominates $\hat{\mathcal{S}}$ whenever each experiment in $\mathcal{S}$ is more informative than the corresponding experiment in $\hat{\mathcal{S}}$.

Let $c(\mathcal{S})$ be the cost associated with any information structure $\mathcal{S}$. The main assumption we make is that the cost c respects the Blackwell order, in the sense we will now make precise.

Assumption 1.1. *If $\mathcal{S}$ (point-wise) Blackwell-dominates $\mathcal{S}'$, then $c(\mathcal{S}) \geq c\left(\mathcal{S}'\right)$.*

Say that an information structure $\mathcal{S}$ is *completely uninformative* if $\mathcal{S}$ is (point-wise) Blackwell-dominated by any $\hat{\mathcal{S}}$. We make the following normalisation.

Assumption 1.2. *If $\mathcal{S}$ is completely uninformative, then $c(\mathcal{S}) = 0$.*

Assumption 1.2 states that acquiring no information entails no costs to the seller.

[10] In principle, the buyer is able to communicate arbitrary messages to the seller. However, the reduction of these messages to type announcements is possible via the revelation principle because the seller has full commitment.

[11] As alluded to in footnote 7, the payoff as written above does not adhere strictly to established notation. However, the meaning should be clear.

1.3.5 The Seller's Problem

Say that the mechanism $(\mathcal{M}, \mathcal{S})$ is *incentive-compatible (IC)* whenever

$$U_i^i(\mathcal{M}, \mathcal{S}) \geq U_i^j(\mathcal{M}, \mathcal{S}) \tag{IC$_i^j$}$$

for all i and j in V. Whenever

$$U_i^i(\mathcal{M}, \mathcal{S}) \geq 0 \tag{IR$_i$}$$

for all i in V, then the mechanism is *individually-rational (IR)*. We say that the mechanism satisfies *limited-liability (LL)* whenever $t^i(s) \in [0, 1]$ for all i and s. Call any mechanism that is incentive-compatible and individually-rational *feasible*.[12]

Let $\mathcal{X}$ be the space of feasible mechanisms. The seller's problem is given by

$$\sup_{(\mathcal{M}, \mathcal{S}) \in \mathcal{X}} R(\mathcal{M}, \mathcal{S}) - c(\mathcal{S}). \tag{OBJ}$$

We will assume that the seller's problem has a solution.

Assumption 1.3. *There exists $(\mathcal{M}, \mathcal{S}) \in \mathcal{X}$ that solves (OBJ).*

Our first proposition states sufficient conditions that guarantee the existence of a solution to the seller's problem.

Proposition 1.1. *Suppose that $\left|S^i\right| \leq N$ for all i for some $N \in \mathbb{N}$ and that $c(\mathcal{S})$ is continuous in $\mathcal{S}$.[13] Then, a solution to (OBJ) exists.*

Proof. Assume the stated conditions. The conclusion follows from the fact that the objective is continuous, and the space of feasible mechanisms is compact. $\square$

Note that our main results do not rely on the antecedents of proposition 1.1. However, proposition 1.1 assures us that our results are not vacuous.

1.4 Main Results

To simplify the later analysis, we first identify a simple class of trading strategies that always contain a solution to (OBJ). We call these trading strategies *extremal*. This enables us to give a characterisation result for the set of optimal mechanisms. In particular, there is always an optimal mechanism that has a simple structure. With an

[12] Recall that a mechanism, by definition, must satisfy limited-liability.

[13] One might be concerned about the need to explicitly specify the topology on the space of information structures. However, the set of information structures is finite-dimensional, and thus has only one natural topology.

arbitrary number of types this simple mechanism employs an extremal trading strategy and an information structure where each experiment is defined on a common signal space containing at most four signals. With two types, there is always an optimal mechanism where the experiment for the high type is degenerate; the trading strategy for the low type is extremal; and, the experiment for the low type consists of at most three signals.

1.4.1 Optimal Trading Strategies

Fix $\mathcal{S}$, and let $\mathcal{M}(\mathcal{S}) = \{\mathcal{M} : (\mathcal{M}, \mathcal{S}) \in \mathcal{X}\}$ be the set of trading strategies that, when paired with $\mathcal{S}$, result in a feasible mechanism. Consider the seller's problem of choosing an optimal trading strategy for a fixed information structure:

$$\max_{\mathcal{M} \in \mathcal{M}(\mathcal{S})} R(\mathcal{M}, \mathcal{S}) . \tag{REV}$$

This is a standard linear program, whose solution has a standard characterisation.

Proposition 1.2. *Suppose that $\mathcal{M}^* = \left\{ \left(x^{i*}, t^{i*} \right) \right\}_{i \in V}$ solves (REV). Then, for each i, there exist non-negative constants $\bar{\xi}^i$ and $\bar{\tau}^i$, along with a collection of weights $\left\{ \lambda^i_j \right\}_{j \in V} \subset \mathbb{R}_+$ such that*

$$t^{i*}(s) \begin{cases} = 1, & \text{if } \sum_{j \neq i} \lambda^i_j \sigma^i_j(s) > \bar{\tau}^i \sigma^i_i(s) \\ \in [0,1], & \text{if } \sum_{j \neq i} \lambda^i_j \sigma^i_j(s) = \bar{\tau}^i \sigma^i_i(s) \\ = 0, & \text{if } \sum_{j \neq i} \lambda^i_j \sigma^i_j(s) < \bar{\tau}^i \sigma^i_i(s) \end{cases}$$

and

$$x^{i*}(s) \begin{cases} = 1, & \text{if } \bar{\xi}^i v_i \sigma^i_i(s) > \sum_{j \neq i} \lambda^i_j v_j \sigma^i_j(s) \\ \in [0,1], & \text{if } \bar{\xi}^i v_i \sigma^i_i(s) = \sum_{j \neq i} \lambda^i_j v_j \sigma^i_j(s) \\ = 0, & \text{if } \bar{\xi}^i v_i \sigma^i_i(s) < \sum_{j \neq i} \lambda^i_j v_j \sigma^i_j(s) \end{cases} .$$

Moreover, $\bar{\xi}^i = \bar{\tau}^i + p_i = \mu_i + \sum_{j \neq i} \lambda^j_i$, and λ^j_i is a Lagrange multiplier for (IC^j_i) while μ_i is a Lagrange multiplier for (IR_i).

Lemma 1.1 below guarantees that it is without loss to assume that the seller always chooses a trading strategy such that $t^i(s) \in \{0,1\}$ and $x^i(s) \in \{0,1\}$ for all $s \in S$ and all $i \in V$. Assuming that the seller choosing a mechanism in this class will be useful when $|V| > 2$.

If, for some i, $\left(x^i, t^i \right)$ satisfies $x^i(s) \in \{0,1\}$ and $t^i(s) \in \{0,1\}$ for all $s \in S^i$, we say that the seller's trading strategy is *extremal for i*. If the seller's trading strategy is extremal for i for all $i \in V$, then we say that the trading strategy is *extremal*. The next result states that any feasible mechanism can always be transformed into another feasible mechanism whose trading strategy is extremal for i. Moreover, this transformation imposes no additional cost to the seller.

Lemma 1.1. *For any feasible mechanism* $(\mathcal{M}, \mathcal{S})$ *where* $t^i(s) \in (0,1)$ *or* $x^i(s) \in (0,1)$ *for some* $s \in S^i$ *and some* $i \in V$*, there exists another feasible mechanism* $\left(\hat{\mathcal{M}}, \hat{\mathcal{S}}\right)$ *such that*

1. $R(\mathcal{M}, \mathcal{S}) = R\left(\hat{\mathcal{M}}, \hat{\mathcal{S}}\right)$;

2. $\hat{t}^i(s) \in \{0,1\}$ *and* $\hat{x}^i(s) \in \{0,1\}$ *for all* $s \in \hat{S}^i$*; and,*

3. $c(\mathcal{S}) \geq c\left(\hat{\mathcal{S}}\right)$.

Proof. Fix a feasible mechanism $(\mathcal{M}, \mathcal{S})$ such that $t^i\left(s'\right) \in (0,1)$ for some i and $s' \in S^i$. We will now construct a new feasible mechanism $\left(\hat{\mathcal{M}}, \hat{\mathcal{S}}\right)$ that satisfies our requirements.

First, let $t^i\left(s'\right) = \alpha \in (0,1)$ and choose $\hat{S}^i = S^i \cup \{s_0, s_1\} \setminus \left\{s'\right\}$ such that $S^i \cap \{s_0, s_1\} = \emptyset$. For each $j \in V$, let

$$\hat{\sigma}^i_j(s) = \begin{cases} \sigma^i_j(s), & \text{if } s \in S^i \setminus \left\{s'\right\} \\ \alpha\sigma^i_j\left(s'\right), & \text{if } s = s_1 \\ (1-\alpha)\sigma^i_j\left(s'\right), & \text{if } s = s_0 \end{cases}.$$

This completes our specification of $\hat{\mathcal{S}}^i$. Since $\hat{\mathcal{S}}^i$ is a garbling of $\mathcal{S}^i$, we must have that $\mathcal{S}^i \succeq \hat{\mathcal{S}}^i$. Let $\hat{\mathcal{S}} = \left\{\mathcal{S}^j\right\}_{j \neq i} \cup \left\{\hat{\mathcal{S}}^i\right\}$. That is, $\hat{\mathcal{S}}$ is the information structure where $\mathcal{S}^i$ is replaced with $\hat{\mathcal{S}}^i$, but is otherwise the same. Notice that $c(\mathcal{S}) \geq c\left(\hat{\mathcal{S}}\right)$.

We now turn to our construction of the trading strategy $\hat{\mathcal{M}}$. Let

$$\hat{t}^i(s) = \begin{cases} t^i(s), & \text{if } s \in S^i \setminus \left\{s'\right\} \\ 1, & \text{if } s = s_1 \\ 0, & \text{if } s = s_0 \end{cases}$$

and

$$\hat{x}^i(s) = \begin{cases} x^i(s), & \text{if } s \in S^i \setminus \left\{s'\right\} \\ x^i\left(s'\right) & \text{if } s \in \{s_0, s_1\} \end{cases}.$$

Otherwise, for $j \neq i$, choose $\left(\hat{x}^j, \hat{t}^j\right) = \left(x^j, t^j\right)$.

We thus have that, for each $j \in V$,

$$\sum_{s \in \hat{S}^i} \hat{\sigma}^i_j(s)\, \hat{t}^i(s) = \sum_{s \in S^i \setminus \{s'\}} \sigma^i_j(s)\, t^i(s) + \alpha\sigma^i_j\left(s'\right) = \sum_{s \in S^i} \sigma^i_j(s)\, t^i(s). \tag{1.4.1}$$

Similarly,

$$\sum_{s\in\hat{S}^i} \hat{\sigma}^i_j \hat{x}^i(s) = \sum_{s\in S^I} \sigma^i_j(s)\, x^i(s) \tag{1.4.2}$$

for each $j \in V$. (1.4.1) and (1.4.2), and the fact that $\left(\hat{x}^j, \hat{t}^j\right) = \left(x^j, t^j\right)$ and $\hat{\mathcal{S}}^j = \mathcal{S}^j$ for $j \neq i$ guarantees that $\left(\hat{\mathcal{M}}, \hat{\mathcal{S}}\right)$ is feasible and $R(\mathcal{M}, \mathcal{S}) = R\left(\hat{\mathcal{M}}, \hat{\mathcal{S}}\right)$.

We can repeat the exercise above until we have that $\hat{t}^i(s) \in \{0,1\}$ for all $s \in \hat{S}^i$. If we also have that $\hat{x}^i(s) \in \{0,1\}$ for all $s \in \hat{S}^i$, then we are done. If not, then a nearly identical procedure as that conducted above will produce a mechanism that satisfies our requirements.[14] □

These two results will be useful will be useful when we characterise the optimal information structure below. Note, however, that while lemma 1.1 gives us a convenient characterisation of the optimal mechanism, we will not always invoke it for all $i \in V$. In particular, when $|V| = 2$, it will be convenient to choose $t^H \in (0,1)$, where H is the high-value buyer.

1.4.2 Optimal Information Structures

We now characterise the optimal information structures in this model. We first tackle the general case, and then consider a setting with two types.

1.4.2.1 General Case

Our first main result is a straightforward consequence of lemma 1.1. The proof is essentially identical to that of the revelation principle.

Theorem 1.1. *There exists an optimal mechanism* $(\mathcal{M}^*, \mathcal{S}^*)$ *satisfying*

1. $x^{*i}(s) \in \{0,1\}$ *and* $t^{*i}(s) \in \{0,1\}$ *for all* $s \in S^{*i}$ *and* $i \in V$*;*
2. $\left|S^{*i}\right| \leq 4$ *for all* $i \in V$*; and,*
3. $S^{*i} = S^{*j}$ *for all* $i, j \in V$*.*

Proof. Take an optimal mechanism $(\mathcal{M}, \mathcal{S})$. Lemma 1.1 guarantees that we can transform this mechanism into one where $x^i(s) \in \{0,1\}$ and $t^i(s) \in \{0,1\}$ for all $s \in S^i$ and $i \in V$. Suppose that $\left|S^i\right| > 4$ for some $i \in V$. The previous transformation guarantees that there must exist a pair $s, s' \in S^i$ such that $x^i(s) = x^i(s')$ and $t^i(s) = t^i(s')$. The signals s and s' can be merged into a single signal s'' without affecting the seller's

[14] To see this, simply exchange the roles of $\hat{x}^i$ and $\hat{t}^i$ in the proof.

revenue or any of the incentive or participation constraints. Moreover, the merging of signals is a garbling of the original information structure, and thus cannot increase the seller's cost. We can repeat this procedure until we have that $\left|S^i\right| \leq 4$ for all $i \in V$. Finally, we can relabel each of the signals so that $S^i = S^j$ for all i and j in V. These transformations give us a mechanism that satisfies our requirements. $\square$

Theorem 1.1 states that it is sufficient for the seller to acquire four signals in any optimal mechanism, regardless of the size of the type space.

1.4.2.2 Two Types

Suppose now that $V = \{v_H, v_L\}$, with $v_H > v_L$. This assumption holds throughout this section. We first state a lemma, whose proof is standard.

Lemma 1.2. *Any optimal mechanism must have that $x^H(s) = 1$ for all $s \in S^H$.*

We now state a proposition asserting that not acquiring information when the buyer announces that their value is v_H is optimal for the seller.

Proposition 1.3. *There exists $\left(\hat{\mathcal{M}}, \hat{\mathcal{S}}\right)$ such that*

1. *$\left(\hat{\mathcal{M}}, \hat{\mathcal{S}}\right)$ solves (OBJ); and,*
2. *$\hat{\mathcal{S}}^H$ is minimal in $\succeq$.*

Proof. Consider a relaxation of (OBJ), where we ignore (IC_L^H). Let $\left(\tilde{\mathcal{M}}, \tilde{\mathcal{S}}\right)$ be a solution to this relaxed problem. Define a new mechanism $\left(\hat{\mathcal{M}}, \hat{\mathcal{S}}\right)$ as follows. Choose $\hat{\mathcal{S}}^H$ to be minimal in $\succeq$, and let

$$\hat{t}^H(s) = \sum_{s' \in S^H} \tilde{\sigma}_H^H(s')\, \tilde{t}^H(s'),$$

for all $s \in \hat{S}^H$. Set $\left(\hat{\mathcal{M}}, \hat{\mathcal{S}}\right)$ be equal to $\left(\tilde{\mathcal{M}}, \tilde{\mathcal{S}}\right)$ otherwise. It is easy to verify that the seller's revenue under $\left(\hat{\mathcal{M}}, \hat{\mathcal{S}}\right)$ is the same as that under $\left(\tilde{\mathcal{M}}, \tilde{\mathcal{S}}\right)$. Moreover, since $\tilde{\mathcal{S}}$ Blackwell-dominates $\hat{\mathcal{S}}$ by construction, $c\left(\tilde{\mathcal{S}}\right) \geq c\left(\hat{\mathcal{S}}\right)$. Finally, since H pays the same transfers in expectation, and we are ignoring (IC_L^H), $\left(\hat{\mathcal{M}}, \hat{\mathcal{S}}\right)$ is feasible in our relaxed problem.[15] Hence, $\left(\hat{\mathcal{M}}, \hat{\mathcal{S}}\right)$ must also solve this relaxed version of (OBJ).

[15] By lemma 1.2, the probability that H is allocated the object is the same across the two mechanisms.

We must now show that $\left(\hat{\mathcal{M}}, \hat{\mathcal{S}}\right)$ solves (OBJ). The minimality of $\hat{\mathcal{S}}^H$ guarantees that[16] $\hat{\sigma}_L^H = \hat{\sigma}_H^H$ which then implies that

$$\sum_{s \in S^H} \hat{\sigma}_L^H(s)\,\hat{t}^H(s) = \sum_{s \in S^H} \hat{\sigma}_H^H(s)\,\hat{t}^H(s).$$

Suppose that $\left(\hat{\mathcal{M}}, \hat{\mathcal{S}}\right)$ violates $\left(IC_L^H\right)$. We then have that

$$U_L^H\left(\hat{\mathcal{M}}, \hat{\mathcal{S}}\right) = \sum_{s \in S^H} \hat{\sigma}_L^H(s)\left[v_L \hat{x}^H(s) - \hat{t}^H(s)\right] > U_L^L\left(\hat{\mathcal{M}}, \hat{\mathcal{S}}\right) \geq 0,$$

so that $\sum_{s \in S^H} \hat{\sigma}_L^H(s)\,\hat{t}^H(s) = \sum_{s \in S^H} \hat{\sigma}_H^H(s)\,\hat{t}^H(s) < v_L$. Thus, the seller achieves a payoff strictly less than what she would achieve had she chosen a completely uninformative information structure, and posted a price equal to v_L. This posted price mechanism is feasible in (OBJ), violating optimality. □

Proposition 1.3 allows us to assume without loss that $S^H = \emptyset$.[17] We assume this in the succeeding discussion. It then follows that the optimal t^H is the largest transfer that is consistent with (IC_H^L) and (IR_H). The argument in the proof above also shows that it is without loss to ignore (IC_L^H), which delivers the following corollary.

Corollary 1.1. *Suppose that $(\mathcal{M}, \mathcal{S})$ solves (OBJ). Then, (IR_L) binds under $(\mathcal{M}, \mathcal{S})$.*

We now turn to a characterisation of the optimal $\left(x^L, t^L\right)$ and $\mathcal{S}^L$. Suppose $(\mathcal{M}, \mathcal{S})$ solves (OBJ) where $\mathcal{S}$ is such that $\left|S^H\right| = 1$. Assume also that $\lambda_H^L > 0$, so that (IC_H^L) binds.[18] Define[19]

$$\bar{\rho}_t = \frac{\bar{\tau}^L}{\lambda_H^L} \text{ and } \bar{\rho}_x = \frac{\bar{\xi}^L v_L}{\lambda_H^L v_H}.$$

We must have, by proposition 1.2 and lemma 1.1, that, for any $s \in S^L$ such that $\sigma_L^L(s) > 0$,

$$t^L(s) \begin{cases} = 1, & \text{if } \dfrac{\sigma_H^L(s)}{\sigma_L^L(s)} > \bar{\rho}_t \\[2ex] \in \{0,1\}, & \text{if } \dfrac{\sigma_H^L(s)}{\sigma_L^L(s)} = \bar{\rho}_t \\[2ex] = 0, & \text{if } \dfrac{\sigma_H^L(s)}{\sigma_L^L(s)} < \bar{\rho}_t \end{cases} \tag{1.4.3}$$

[16] Suppose that $\hat{\sigma}_L^H \neq \hat{\sigma}_H^H$ and let $\mathcal{S}^H$ be the uniform distribution on S. We would then have that $\hat{\mathcal{S}}^H \succ \mathcal{S}^H$, violating the minimality of $\hat{\mathcal{S}}^H$.

[17] This is why it is inconvenient to use lemma 1.1 for H. We can insist that $t^H \in \{0,1\}$, but then we must have that $\left|S^H\right| > 1$ in order to satisfy feasibility and optimality. However, it is simpler to fix $S^H = \emptyset$, so we forego lemma 1.1 here.

[18] Recall that λ_H^L is a Lagrange multiplier for (IC_H^L).

[19] $\bar{\tau}^L$ and $\bar{\xi}^L$ are defined in proposition 1.2.

and

$$x^L(s)\begin{cases} =1, & \text{if } \bar{\rho}_x > \dfrac{\sigma_H^L(s)}{\sigma_L^L(s)} \\[2ex] \in\{0,1\}, & \text{if } \bar{\rho}_x = \dfrac{\sigma_H^L(s)}{\sigma_L^L(s)} \\[2ex] =0, & \text{if } \bar{\rho}_x < \dfrac{\sigma_H^L(s)}{\sigma_L^L(s)} \end{cases}. \tag{1.4.4}$$

Moreover, it is easy to see that for any $s \in S^i$ such that $\sigma_L^L(s) = 0$, it is optimal to set $t^L(s) = 1$ and $x^L(s) = 0$. Hence, we can take $\sigma_H^L(s)/\sigma_L^L(s)$ to be ∞ when $\sigma_L^L(s) = 0$. We now have the following result.

Proposition 1.4. *For any optimal mechanism such that (IC_H^L) binds, there exists another optimal mechanism where either*

$$\left(x^L, t^L\right) \in \left\{(1,0),(0,0),(0,1)\right\},$$

or

$$\left(x^L, t^L\right) \in \left\{(1,0),(1,1),(0,1)\right\}.$$

Proof. First suppose that $\lambda_H^L > 0$. The result then follows from (1.4.3) and (1.4.4).

Suppose instead that $\lambda_H^L = 0$. Using notation from proposition 1.2, and the fact that we can ignore (IC_L^H), we have that $\bar{\tau}^L = \mu_L - p_L$. If $\bar{\tau}^L < 0$, we would have that $t^L(s) = 1$ for all S^L, which violates (IR_L). Thus, $\bar{\tau}^L \geq 0$, so that $\bar{\xi}^L = \mu_L \geq p_L > 0$. Proposition 1.2 then implies that $x^L(s) = 0$ for all $s \in S^L$ such that $\sigma_L^L(s) > 0$. (IR_L) then enforces that $t^L(s) = 0$ for all $s \in S^L$ such that $\sigma_L^L(s) > 0$. It is clearly optimal to set $x^L(s) = 0$ and $t^L(s) = 1$ for any $s \in S^L$ such that $\sigma_L^L(s) = 0$.[20] This completes the proof. □

We can also state a corresponding result for the case where (IC_H^L) does not bind.

Proposition 1.5. *For any optimal mechanism where (IC_H^L) does not bind, there exists another optimal mechanism where we have that $x^H = x^L = 1$, $t^H = v_H$, and $t^L \in \{0,1\}$.*

Proof. Consider an optimal mechanism, and suppose that (IC_H^L) does not bind. Lemma 1.2 implies that $x^H = 1$. Optimality then requires that $t^H = v_H$.[21] Suppose that $x^L(s) < 1$ for some $s \in S^L$. We can then increase $x^L(s)$ by ε and increase $t^L(s)$ by $v_L\varepsilon$, which keeps L indifferent but increases revenue. We can also choose ε small enough to respect (IC_H^L). This is a violation of our assumption of optimality. □

[20] Note that if the $\left(x^L, t^L\right)$ that we constructed results in (IC_H^L) to be slack, then the mechanism cannot be optimal.

[21] Recall that we assume that $S^H = \emptyset$. Hence, if $t^H \neq v_H$, then either (IR_H) is violated, or we can increase t^H without violating any constraints. Both these cases are impossible.

We are now ready to state our second main result, which specialises theorem 1.1 to the binary-type case.

Theorem 1.2. *There exists an optimal mechanism* $(\mathcal{M}^*, \mathcal{S}^*)$ *satisfying*

1. $\left|S^H\right| = 1$ *and* $\left|S^L\right| \leq 3$*; and,*

2. $x^H = 1$, $x^L(s) \in \{0,1\}$, *and* $t^L(s) \in \{0,1\}$ *for all* $s \in S^L$.

Proof. The fact that $\left|S^H\right| = 1$ and $x^H = 1$ follows from 1.3 and 1.2. The rest of the result is deduced from propositions 1.4 and 1.5, along with the merging argument in the proof of theorem 1.1. □

1.5 Discussion

1.5.1 Limited Liability

Throughout our analysis, we assumed that $t^i(s) \in [0,1]$ for all $s \in S^i$ and $i \in V$. This is equivalent to assuming a strong form of limited liability, since both the buyer and the seller face a limited-liability constraint. It turns out that if we do not assume limited liability, under mild assumptions on the cost function c, we cannot guarantee the existence of a solution to the seller's problem. This is because of a discontinuity in the seller's objective when we weaken our limited-liability assumption. Furthermore, whenever a solution to the seller's problem exists under no limited liability, this solution must involve the full extraction of surplus at zero cost.

Assumption 1.4. *Let* $\mathcal{S}_0$ *be a completely uninformative information structure. Then,* $c(\mathcal{S}) \to c(\mathcal{S}_0) = 0$ *as* $\mathcal{S} \to \mathcal{S}_0$.

We say that the cost function c is *continuous at zero* whenever c satisfies assumption 1.4. Moreover, say that c is *strictly increasing at zero* if $c(\mathcal{S}) > 0$ whenever $\mathcal{S} \succ \mathcal{S}_0$, where $\mathcal{S}_0$ is completely uninformative.

Theorem 1.3. *Suppose that c is continuous at zero, and that there is no limited liability, so that* $t^i(s) \in \mathbb{R}$ *for all* $s \in S^i$ *and all* $i \in V$*. Then, for every* $\varepsilon > 0$*, there exists a mechanism* $(\mathcal{M}, \mathcal{S})$ *such that*

$$\left| R(\mathcal{M}, \mathcal{S}) - c(\mathcal{S}) - \sum_{i \in V} p_i v_i \right| < \varepsilon.$$

Thus, when c is strictly increasing at zero, the seller's objective (OBJ) is discontinuous at any completely uninformative information structure. The same conclusion holds if instead we assumed that $t^i(s) \leq 1$ *for all* $s \in S^i$ *and all* $i \in V$.

The statement of proposition 1.3 for the case with no limited liability is actually a special case of the result of Cremer and McLean (1985; 1988).[22]

Proof. We will prove the proposition for the case where $|V| = 2$. The general case can be handled similarly; the only additional difficult is added notation.

It is enough to prove the result for the case where $t^i(s) \leq 1$ for all $s \in S^i$ and $i \in V$. Consider a mechanism where $S^H = \emptyset$, $S^L = \{s_H, s_L\}$, $x^H = x^L(s) = 1$ for $s \in S^L$, $t^H = v_H$, and $t^L(s_H) = 1$. We will suppress the superscript in $\sigma_i^L(s)$, because it is clear that this probability must refer to the experiment for the low type. We will construct a mechanism that extracts full surplus. This mechanism must satisfy

$$\begin{aligned}\sigma_H(s_H) + \sigma_H(s_L)\, t^L(s_L) &\geq v_H \\ \sigma_L(s_H) + \sigma_L(s_L)\, t^L(s_L) &= v_L.\end{aligned}$$

The inequality corresponds to (IC_H^L) while the equality corresponds to (IR_L). Notice that (IR_H) and (IC_L^H) are satisfied by construction. Define[23], for each $x > 0$,

$$\begin{aligned}\sigma_H(s_L) &= \frac{1}{x} \\ \sigma_L(s_L) &= \frac{1}{\log x} \\ t^L(s_L) &= (v_L - 1)\log x + 1.\end{aligned}$$

By construction, (IR_L) is satisfied. We claim that for large enough x, (IC_H^L) is also satisfied. Moreover, $c(\mathcal{S})$ can be made arbitrarily close to zero by choosing x large enough. First, observe that

$$\sigma_H(s_H) + \sigma_H(s_L)\, t^L(s_L) = 1 + (v_L - 1)\frac{\log x}{x}.$$

Since

$$\lim_{x\to\infty} \frac{\log x}{x} = 0,$$

our first assertion about (IC_H^L) follows. Moreover, since $\sigma_H(s_H) \to 1$ and $\sigma_L(s_H) \to 1$ as $x \to \infty$, the information structure we constructed is completely uninformative in the limit. This completes the proof. □

[22] See also McAfee and Reny (1992).

[23] The functions x^{-1} and $(\log x)^{-1}$ are not unique. Any pair of functions f and g that converge to zero and satisfy

$$\lim_{x\to\infty} \frac{f(x)}{g(x)} = 0$$

will do.

1.5.2 Tightness of Bounds on $\left|S^i\right|$

Theorems 1.1 and 1.2 construct an upper bound on the complexity of a simple optimal mechanism. It is natural to ask whether this bound is tight. Could it be optimal for the seller to choose an even simpler mechanism? In general, the answer is no, at least in the binary-type case. This is a consequence of the next result.

Theorem 1.4. *Suppose that $|V| = 2$. There exists a cost function c and primitives p, v_L, and v_H such that any mechanism with $\left|S^L\right| < 3$ is not optimal.*

Proof. Choose p, v_L, and v_H along with an information structure $\mathcal{S}$ such that the unique solution to (REV) given $\mathcal{S}$ is of the form given in 1.4.[24] Let c be given by

$$c\left(\mathcal{S}'\right) = \begin{cases} 1, & \text{if } \mathcal{S} \not\succeq \mathcal{S}' \\ 0, & \text{if } \mathcal{S} \succeq \mathcal{S}' \end{cases}.$$

It is easy to see that for such a cost function, there does not exist an optimal mechanism with $\left|S^L\right| < 3$. □

To complete the proof of the previous result, we give the following example.

Example 1.1. Let $p_L = p_H = 1/2$, with $v_L = 1/3$ and $v_H = 2/3$. Suppose that $S^H = \emptyset$, and that $S^L = \{s_1, s_2, s_3\}$, with

$$\begin{pmatrix} \sigma_H(s_1) & \sigma_H(s_2) & \sigma_H(s_3) \\ \sigma_L(s_1) & \sigma_L(s_2) & \sigma_L(s_3) \end{pmatrix} = \begin{pmatrix} 0.57 & 0.24 & 0.19 \\ 0.56 & 0.24 & 0.2 \end{pmatrix}.$$

The optimal trading strategy is given by $x^H = 1$, $t^H = 37/75$, and

$$x^L(s) = \begin{cases} 1, & \text{if } s = s_1 \\ 0, & \text{if } s \neq s_1 \end{cases}$$

$$t^L(s) = \begin{cases} 1, & \text{if } s = s_3 \\ 0, & \text{if } s \neq s_3 \end{cases}.$$

1.6 Conclusion

In this paper, we studied a model of screening with information acquisition. We found that there always exist simple optimal mechanisms. Without limited liability on both

[24] The example given below demonstrates this is possible.

the buyer and the seller, the seller can always extract full surplus at vanishingly small cost.

Our results heavily exploit the linear structure of the screening problem. It would be useful to know whether there is a similar characterisation of the optimal information structure in a version of this model where buyers have concave utility. Moreover, it would also be worthwhile to explore more specific information acquisition cost functions within the class that respects the Blackwell order.[25] We leave this to future work.

[25] One example is the entropy cost function used in the rational inattention literature. Numerical computations suggest that for this cost function, the seller cannot do better than the optimal two-signal information structure.

2 Markovian Information Design in Games of Strategic Experimentation

2.1 Introduction

Internet platforms such as Google, Yelp, TripAdvisor, and even Facebook, are some of our most ubiquitous sources of information. However, much of the information they provide is generated by their users, and not something they produce themselves. Moreover, any information these Internet platforms reveal affects their users' subsequent choices which, in turn, changes the information that is fed back to the platform. In other words, these Internet platforms face an information disclosure problem where their choices impact the production of information. The information they reveal today affects the information they can reveal tomorrow.

Consider the example of TripAdvisor. If TripAdvisor posted a large number of positive reviews of a hotel in a short period of time, then its users have little incentive to experiment with other hotels that have few or no reviews. This is inefficient because it may well be the case that one of the latter hotels is much better than the former, but we are never able to discover this fact due to inefficient underexperimentation. Trivially, it would also be inefficient for TripAdvisor to never post reviews, because this leads to the underutilisation of information that they have. Hence, the timing of information revelation is of natural interest in this setting. If Internet platforms cared only about efficiency[1], when should they disclose information?

Suppose instead that an Internet platform has some objective other than efficiency, and we are concerned that it is distorting information to achieve its own, potentially socially undesirable, ends. For example, people have expressed the criticism, made particularly alarming by their monopoly, that Google opaquely manipulates the information they provide their users.[2] This complaint has also been levelled against other platforms such as Yelp and Facebook.[3] This leads us to consider the content of information disclosure. Under what conditions must an Internet platform reveal information transparently?

[1] More precisely, suppose that they wish to maximise the sum of their users' payoffs.

[2] On 27 June 2017, the European Commission fined Google €2.42 billion for breaching antitrust rules by manipulating search results. (`http://bit.ly/EUGoogFine`)

[3] A petition on change.org to stop Yelp from censoring reviews can be found at `http://bit.ly/ChangeYelp`. Former Facebook workers have admitted to suppressing conservative news on its 'trending' news section. (`http://bit.ly/FacebookSuppress`)

To address the questions posited above, we embed the canonical model of strategic experimentation with exponential bandits (Keller et al., 2005) into an information design setting. The Internet platform is the designer, and their users are agents who have a risky arm and a safe arm. At each instant of time, agents allocate shares of a perfectly divisible unit resource to each arm to maximise their expected discounted payoffs. The safe arm gives known flow payoffs while the risky arm produces (known, lump-sum) payoffs (called successes) at some unknown rate. Use of the risky arm when it is not known to be better than the safe arm is experimentation. A good risky arm yields payoffs at a rate more profitable than the safe arm; a bad risky arm never generates any payoffs. All agents have risky arms of the same type, and externalities are informational only. The classic setting that Keller, Rady, and Cripps study is a game of perfect monitoring: agents always observe each others' actions and payoffs. Here, the only information that agents obtain about other agents comes exclusively from what the designer discloses to them. More specifically, at the start of the game, the designer commits to an information structure—sometimes also called a mechanism[4]—that maps histories (of actions and payoffs) into (for now, private) signals sent to each agent. Thus, our designer observes all agents' histories, but she is equally unaware about the common type of the risky arms.[5,6]

We shall restrict attention to a tractable class of information structures that possess the property that any signal they transmit must depend only on a simple summary statistic of the history of the game: the designer's posterior over the type of the risky arm. These shall be called Markovian information structures, or Markov mechanisms for short. We study the equilibria[7] these information structures induce. We show, in the benchmark setting with deterministic mechanisms[8], that all such equilibria possess the common feature that the designer discloses no information, or all (payoff-relevant) information, individually to each agent. Partial information disclosure is impossible. Differential information disclosure across different agents, however, is available to the designer. Naturally, partial information disclosure is possible with, and only with, stochastic mechanisms.

We then obtain a starker characterisation of equilibria by focusing on those that can be implemented with symmetric monotone *recommendation* mechanisms. These mechanisms provide a simple generalisation of symmetric Markovian perfect-monitoring equilibria. A recommendation mechanism is symmetric when the designer, for any given

[4]This terminology is appropriate because information design is in some sense dual to mechanism design. See Bergemann and Morris (2016b, 2017).

[5]In light of the motivation above, we will not explicitly specify the designer's preferences. Note that the information structure is common knowledge, and so, once one has been fixed, the designer's objective plays no further role. Our focus will therefore be on implementation.

[6]We abstract from the problem of eliciting truthful reports from agents about their experience. This is without loss for any equilibrium we construct.

[7]We look for weak perfect Bayesian equilibria in Markov strategies.

[8]The real-world analogue of a mechanism is an Internet platform's algorithm, which is deterministic.

history[9], sends the same recommendation to all agents.[10,11] Monotonicity means that recommendations must be monotonic in the designer's posterior. In any equilibrium where agents obey the designer, each symmetric monotone (recommendation) mechanism can be identified with a cut-off on the designer's posterior. Above this cut-off, all agents use the risky arm exclusively, and the designer discloses no information. Below it, the designer reveals all information and the agents play the unique symmetric Markov-perfect equilibrium (MPE) of the perfect monitoring game. Thus, symmetric monotone recommendation mechanisms correspond to information structures that completely withhold information until aggregate experimentation reaches some threshhold, at which point all information is then released.[12]

A symmetric monotone mechanism can be identified with a cut-off. However, not all cut-offs induce equilibria where agents obey the designer's recommendations. Hence, we characterise the cut-offs—called implementable—that do, and give necessary and sufficient conditions for a cut-off to belong in the implementable set. When agents are patient enough, the implementable set forms a closed (but possibly degenerate) interval.[13] Furthermore, we can use the aforementioned necessary and sufficient conditions to study the efficiency and transparency properties of these mechanisms. We find that for each pessimistic enough prior[14], there is a large enough number of agents such that the set is a singleton corresponding to the perfect monitoring outcome.[15] That is, symmetric monotone mechanisms are maximally transparent under these conditions. This is because the Markov property of the mechanism limits the designer's ability to selectively reveal information in a way that prevents free-riding, and the temptation to free-ride on others' experimentation is greatest when there are a large number of agents. More explicitly, under these conditions, an agent's contribution to speeding up the disclosure of information via inducing decay in the designer's posterior (conditional on the absence of success) is small, and therefore cannot outweigh the (net) expected marginal cost of experimentation. Consequently, the designer cannot induce full experimentation. On the other hand, when agents begin the game with a prior optimistic enough about the risky arm, then the implementable set is large enough to contain cut-offs inducing out-

[9] Note that these histories need not be on the equilibrium path.

[10] An obedient equilibrium strategy induced by a symmetric recommendation mechanism satisfies a property called strong symmetry. (A strategy is obedient whenever the actions it specifies always coincide with the recommendations.) Abreu (1986) and Abreu et al. (1986) study strongly symmetric equilibria (SSE) in repeated Cournot games. Cronshaw and Luenberger (1994) give a characterisation of these equilibria in general repeated games with perfect monitoring. Hörner et al. (2015) also use strong symmetry to generalise symmetric Markovian perfect-monitoring equilibria. However, their work uses SSE to generalise the Markov property of (perfect monitoring) equilibria, whereas we maintain the Markov property and instead use the implied strong symmetry to obtain a simple generalisation to perfect monitoring.

[11] Symmetric recommendation mechanisms correspond to public information structures, where the same signal is always sent to all agents.

[12] This is reminiscent of the publication of reviews on the Apple App Store, where app reviews are hidden until an app has some required number of reviews.

[13] This is not a limit result. We only require that the discount rate be close enough to zero.

[14] Any prior for which a myopic agent would be unwilling to use the risky arm will do. This condition can be interpreted as experimentation being 'costly'.

[15] The implementable set shrinks to a singleton for a finite but large number of agents.

comes that are always more efficient than any outcome under perfect monitoring. The reason for this is, when agents are already optimistic, it becomes easy for the designer to convince agents to experiment without having to reveal to them any information. The designer thus delays the disclosure of information for as long as possible, at least until the time when it might be efficient for all agents to switch to the safe arm.[16] This means that there is often a direct trade-off between transparency and efficiency. An efficient mechanism must, to some extent, be opaque.

The benchmark setting described above considers the case with *conclusive good news*. The arrival of payoffs from the risky arm immediately reveals, at least to the designer and the agent who enjoys the payoff, that the risky arm is good. We can instead consider a variation of the model with *conclusive bad news* (Keller and Rady, 2015). In this setting, both arms incur *costs*.[17] A good risky arm never generates any costs, while a bad risky arm generates (lump-sum) costs (failures) at a rate higher than the flow cost of using the safe arm. With conclusive bad news, barring trivial cases, the equilibrium induced by the class of symmetric monotone mechanisms is unique and coincides with the unique symmetric MPE under perfect monitoring. Moreover, within this class of mechanisms, there is no trade-off between transparency and efficiency. This is because, under some conditions on the prior, the only other equilibria the designer can induce in this class entail no experimentation, even when some would be efficient.

These results are relevant to the regulation of Internet platforms, where society may wish to maximise both transparency and efficiency. Our model suggests that one needs to think about the kind of information these platforms provide. If this information corresponds to that generated by a bad news bandit, then these platforms may have no choice but to be transparent, at least to the extent that their algorithms resemble our symmetric monotone mechanisms. If not, then perhaps their algorithms ought be regulated so that they are symmetric and monotone. Further, if their information corresponds to a good news bandit, then these results suggest that having a large user base is exactly what enforces transparency. This means that if one values transparency, Internet platform monopolies may not be undesirable. More generally, transparency is best served when the designer is limited to mechanisms that inhibit her ability to prevent free-riding. However, if what one really cares about is efficiency, then criticisms over lack of transparency in this case, unlike that of bad news, may be misguided.

In the course of our analysis of symmetric monotone mechanisms under good news, we also uncover a novel kind of encouragement effect, as initially identified in the literature on strategic experimentation (Bolton and Harris, 1999; Keller et al., 2005; Keller and Rady, 2010, 2015). In the perfect monitoring case, the encouragement effect captures the motive of agents to experiment more than they would individually because

[16] Note, however, that implementing this mechanism involves more than just delaying information disclosure until some fixed time. Information is revealed when the designer's posterior crosses a certain threshold, but the time that this occurs is a function of the (equilibrium) amount of experimentation. Treating the time of disclosure as fixed will not be enough to belay free-riding.

[17] See Keller and Rady (2015) for examples of situations this version of the model can be applied to.

experimentation subsequently causes other agents to do the same, generating useful information. Our encouragement effect is different. First, we find it in a setting with conclusive good news, under which there is no encouragement effect with perfect monitoring. Second, the encouragement effect in the perfect monitoring case hinges on the incentives presented by information generated in the future. The encouragement effect here is about accelerating the disclosure of information generated in the past. Finally, in the perfect monitoring case, larger numbers of agents typically strengthen the encouragement effect. We find the opposite here.

The Markov assumption is clearly with loss of generality. For instance, it rules out the use of punishments via the denial of information, which can be used to support a wider range of equilibrium outcomes.[18] However, this assumption is not just for the sake of tractability; it has economic content as well. The Internet platforms that we model as information designers probably do not have the ability to deny their users information in order to bring about outcomes they desire. Even if they did have this ability, there are perhaps market-related reasons that prevent them from doing so, and we can view our assumption as a reduced-form way of capturing this limitation. Moreover, given the vast quantities of data that these platforms manage, it seems likely that the information they provide their users does in fact depend only on simple summary statistics like our designer's posterior.[19]

While not immediately apparent, the restriction to recommendation mechanisms, when coupled with the Markov requirement, is also with loss. In particular, there exist public Markovian information structures yielding symmetric equilibria that cannot always be replicated using symmetric Markov recommendation mechanisms. One example is the information structure in Bonatti and Hörner (2011), where actions are private and successes are public.[20] In this case, there is a unique symmetric equilibrium where agents use the risky arm exclusively until the common (equilibrium) posterior belief reaches the level where an agent experimenting in complete isolation would have stopped, at which point they all switch to the safe arm.[21] Our results detailed above show that for pessimistic enough priors and a large enough number of agents, the equilibrium in this setting is not implementable with symmetric Markov recommendations. The difference is driven by out-of-equilibrium behaviour. With private actions and public successes, deviations from the risky arm are not particularly attractive because other agents never observe these deviations, and hence never respond to them. With Markov recommendations, the same deviations slow the decay in the designer's posterior, which leads to all other agents experimenting for longer than they would have otherwise. This makes these deviations attractive.

Other applications of our model include information sharing in organisations. For ex-

[18] In particular, the first-best is *always* attainable when the designer uses grim-trigger type punishment strategies. See section 2.6.1.

[19] Consider, for example, user review scores, or Google's PageRank algorithm.

[20] Bonatti and Hörner (2011) also have payoff externalities, which we do not.

[21] See Bonatti and Hörner (2017b), footnote 28.

ample, managers of separate R&D teams may face the problem of information disclosure across teams that work on similar projects with uncertain prospects of success. However, applications aside, this model is arguably of significant theoretical interest as well. In the same way that zero-sum games form a useful benchmark for more general games by highlighting their adversarial aspects, this model, as it abstracts away from payoff externalities, brings to the fore pure informational issues in dynamic information design. In particular, our results exhibit a class of repeated games with learning and imperfect public monitoring where efficiency can be attained with limited resort to rewards and punishments.[22] It would be useful to know whether this is an artefact of the assumptions we have made, or if this result emerges in more general games. We leave this question to future research.

2.2 Literature Review

This chapter is most closely related to two major strands of literature: experimentation with bandits, and information design.

There is an expansive literature on bandit problems.[23] Rothschild (1974) is commonly credited for introducing them to economics. The two-armed Poisson bandit[24] we use here is due to Presman (1991), and their application to strategic settings of perfect monitoring was studied in Keller et al. (2005), who consider a setting with conclusive good news; Keller and Rady (2010), inconclusive good news; and Keller and Rady (2015), both conclusive and inconclusive bad news. However, the first use of bandits in a game of strategic experimentation was by Bolton and Harris (1999), who endowed players with Brownian bandits. Technical differences notwithstanding, the equilibria in these games all exhibit inefficient under-experimentation, and, with the exception of Keller et al. (2005), an encouragement effect.

Poisson bandits, and their discrete-time analogues, have also been applied to a rich set of strategic environments different from, but related to, those already mentioned.[25] Two closely related papers are by Bonatti and Hörner (2011, 2017b). Bonatti and Hörner (2011) study a setting with public conclusive good news but private actions. They find that there is still inefficient under-experimentation, but less so than under perfect monitoring. In Bonatti and Hörner (2017b), also with private actions but public conclusive bad news, experimentation is inefficient, and even more so than under perfect monitoring. These results are similar to some of our results on efficiency, but their

[22]See Sugaya and Yamamoto (2015) for a relevant folk theorem.

[23]For comprehensive introductions, see Berry and Fristedt (1985) or Gittins et al. (2011). Bergemann and Välimäki (2008) give a short overview with a brief survey of economic applications.

[24]Exponential bandits are a special case of Poisson bandits.

[25]See, for instance, Bergemann and Hege (2005); Strulovici (2010); Klein and Rady (2011); Hörner and Samuelson (2013); Cripps and Thomas (2016); Di Pei (2016); Guo (2016); Guo and Roesler (2016); Halac et al. (2016b); Bonatti and Hörner (2017a). Hörner and Skrzypacz (2016) give an excellent survey of this literature.

focus on design issues is different from ours. In particular, they treat the information structure as fixed whereas we focus on treating it as a design variable.[26] Another pair of closely related papers are Hörner et al. (2015) and Heidhues et al. (2015), who study the construction of efficient equilibria. Hörner et al. (2015) show that efficient payoffs are achievable in strongly symmetric equilibria under perfect monitoring with inconclusive good news, but are not attainable in such equilibria when news is conclusive. Heidhues et al. (2015) show that efficiency can be attained in the conclusive case in sequential equilibria with communication and private payoffs. The equilibria these two papers construct rely on the use of punishments to deter deviations; we do not have access to these here. Dong (2016) studies a setting with perfect monitoring and conclusive good news, as in Keller et al. (2005), but one of the players has private information about the distribution of payoffs from the risky arm. If our designer could reveal information about the state instead of choosing the monitoring structure, then the outcomes she could induce would resemble the equilibria constructed by Dong.

In information design[27], the seminal paper is Kamenica and Gentzkow (2011), who consider a general static problem with a single agent. Taneva (2016) and Mathevet et al. (2017) extend the analysis to static games while Ely (2017) and Ball (2017) extend the single agent problem to a dynamic setting.[28] More substantively, Ely's and Ball's analyses focus on the case where the underlying information process is exogenous whereas we study the problem of revealing information that is endogenously generated. Kremer et al. (2014) and Che and Hörner (2017) are motivated by many of the same issues we are, and consider a design problem reminiscent of the one we have here except that agents are myopic and arrive sequentially.[29] A setting with information disclosure to non-myopic agents is of natural theoretical interest, but is of interest for the purposes of applications as well: Google's users, for example, typically use Google's services repeatedly, so it is worth considering a model where the designer and agents interact dynamically, as they do here. In particular, with myopic agents, the designer is not constrained by an agent's desire to free-ride. One noteworthy resulting difference is that Kremer, et al. find that a benevolent designer can achieve the first-best when the number of agents is large whereas we have the exact opposite result, at least within the class of information structures we consider. Smolin (2015) studies a setting similar to ours with a single agent who does not observe their own payoffs, and is thus motivated instead by issues of information disclosure in the presence of disagreement between the

[26] A collection of other papers (Rosenberg et al., 2007, 2013; Murto and Välimäki, 2011, 2013) also consider a variety of information structures for games of strategic experimentation (typically with irreversibility of switching between arms), but none of them directly allow the information structure to be a subject of design. Halac et al. (2016a) consider the disclosure problem in a similar setting to ours, but they wish to characterise optimal contests for a single success among agents who do not intrinsically care about the outcomes of the risky arm.

[27] See Bergemann and Morris (2016b, 2017) for surveys.

[28] Ely (2017) also considers a stylised two-player coordination game commonly used to model bank runs.

[29] To the best of my knowledge, this is the first paper that explicitly considers the information design problem in a setting of strategic experimentation with long-lived agents. Aside from the pair named above, there are a number of other papers examining the problem with short-lived agents such as Mansour et al. (2015) and Papanastasiou et al. (2017).

designer and the agent over optimal allocation between the risky and safe arms.

This chapter is also related to the literature on repeated games. As mentioned, our model formally corresponds to to a repeated game with learning and imperfect monitoring. Sugaya and Yamamoto (2015) prove a folk theorem in this class of environments.[30] As is typical in this literature, their folk theorem relies on constructing complicated strategies to achieve efficiency. The information structure and strategies we construct are comparatively simple. Our results contrasting the good news and bad news environments are reminiscent of the results in Abreu et al. (1991), who find that efficiency is achievable with bad news but not with good news. However, their model is entirely different. News in their model is about deviations from equilibria; ours is about the type of the risky arms. Sugaya and Wolitzky (2017) call our information structure the *universal monitoring structure*, which they use to derive bounds on the sequential equilibrium payoff set in repeated games with imperfect private monitoring. Finally, the equilibria we construct can be viewed as a type of mediated equilibrium. Rahman (2014) uses mediated equilibria to construct collusive outcomes in a repeated Cournot game.

Our information designer is also very similar to the mediator in Myerson (1986). One important difference is that Myerson's mediator does not observe her agents' information, but must ask them to report this truthfully. The requirement that agents report their information typically constraints the mediator in a way our designer seemingly is not. However, with conclusive news, the outcome of any implementable symmetric monotone recommendation mechanism is also an equilibrium in the game with the mediator who does not directly observe agents' information. This is because the designer can always ask agents to report that they have observed a success. When news is conclusive, agents do not have the incentive to withhold this information.

2.3 Model

This section lays out the setup of the model. We first describe the basic game and give the definition of an information structure.[31] We then move on to describe the designer's belief dynamics, followed by a discussion of strategies and equilibrium. We use the following convention for indices: i and j refer to agents; D refers to the designer; and, t and τ refer to time. Subscripts will be dropped when the interpretation is clear from the context. The partial derivative of a function f with respect to x is denoted by $\partial_x f$. All spaces and functions are taken to be measurable. A discussion of many of the assumptions made below can be found in section 2.6 while omitted derivations and proofs are in appendix 2.9.

[30] See also Wiseman (2005, 2012); Fudenberg and Yamamoto (2010, 2011); Yamamoto (2014).

[31] Decomposing a game into a basic game and an information structure is standard. See Bergemann and Morris (2016a).

2.3.1 Basic Game: Players, Actions, and Payoffs

There are $N > 1$ agents (they[32]) and a designer (she). We will commit the usual notional abuse of using N to refer to both the set of agents and the number of agents. Time $t \in [0, \infty)$ is continuous and the horizon is infinite. All agents have a common discount rate of $r > 0$. There is a state of the world $\theta \in \Theta = \{0, 1\}$ that is is unknown to both the designer and all agents. $p_0 = \Pr(\theta = 1) \in (0, 1)$ is the common prior over the state of the world. Agent i's posterior belief[33] at the start of time t is p_{it}; the designer's is p_{Dt}.

Agents each face a two-armed Poisson bandit consisting of a risky arm R and a safe arm S. At time t, each agent $i \in N$ chooses a fraction $k_{it} \in [0, 1]$ of a perfectly divisible resource to allocate to R. The remainder of the resource is allocated to S. Each risky arm generates lump-sum payoffs h at a Poisson rate $\lambda_\theta k_{it}$ that depends on the state. Consequently, agents' actions control the rates of conditionally independent Poisson processes $\left\{\nu_{it}^{\theta,k} : t \geq 0, i \in N\right\}$ where $\nu_{it}^{\theta,k}$ is the number of lump-sum payoffs observed by agent i up to (the end of) time t. We assume that $\lambda_1 > \lambda_0 = 0$. Each safe arm generates a known flow payoff s for each unit of the resource allocated to it. That is, when agent i allocates k_{it} to R at time t, they receive a flow payoff of $(1 - k_{it})\, s$. We say that agent i experiments at time t when $k_{it} > 0$. Hence, given an (integrable) experimentation path $\{k_{it} : t \geq 0\}$, and a realisation of the process $\left\{\nu_{it}^{\theta,k} : t \geq 0\right\}$, agent i's realised payoff is

$$\int_0^\infty re^{-rt} \left(h\, \mathrm{d}\nu_{it}^{\theta,k} + (1 - k_{it})\, s\, \mathrm{d}t\right).$$

Agent j's actions do not enter directly into i's payoff when $i \neq j$, so this is a game of informational externalities only. Call the arrival of a lump-sum payoff, if observed, the arrival of *news*.[34]

We initially assume that $\lambda_1 h > s > \lambda_0 h = 0$. This means that agents strictly prefer R when they know that $\theta = 1$, and they strictly prefer S when $\theta = 0$. Under this assumption, we call the risky arm good when $\theta = 1$ and bad when $\theta = 0$. We shall refer to this assumption as the *good news* case. Naturally, observing the arrival of the lump sum h is good news: agents enjoy their highest feasible payoff, and they revise upward their belief that the risky arm is good. Later, we shall assume that $\lambda_1 h < s < \lambda_0 h = 0$. This is the *bad news* case.[35] Agents strictly prefer R when they know that $\theta = 0$; they strictly prefer S when they know that $\theta = 1$. With bad news, we call the risky arm

[32] We will sometimes abuse our linguistic convention by also referring to a single agent as 'they'. No confusion should result.

[33] Throughout the chapter, we model beliefs as left continuous processes, so that they describe beliefs at the start of time t before any further information is revealed. This allows us to use beliefs as state variables directly instead of working with their left-continuous versions.

[34] This terminology, while convenient, is somewhat misleading. The absence of any lump sum payoff arrivals can also be thought of as news since agents will update their posterior if they knew this. However, we this caveat in mind, we shall continue to use the term news to refer to the arrival of lump-sum payoffs.

[35] All other cases are trivial. To see this, suppose that $s \geq \lambda_1 h > \lambda_0 h$. It is then never optimal to use the risky arm. Other cases are handled similarly.

bad when $\theta = 1$ and good when $\theta = 0$. Under both assumptions, since $\lambda_0 = 0$, news is *conclusive*. This means that agents infer with certainty that $\theta = 1$ upon observing the arrival of a lump-sum payoff. In the good news case, the arrival of news will also be called a success, whereas in the bad news case, the arrival of news is a failure.

The designer chooses the information structure of this game under perfect commitment. We will be more precise about what this means in the next subsection. We wish to remain agnostic about the designer's preferences, so, for most of our analysis, these preferences shall remain unspecified. This means that the focus of this model will be on *implementation*. In particular, we are interested in the outcomes that the designer can implement given her choice of information structure for this game. However, in a later section, we shall also briefly examine some natural specifications for the designer's preferences.

2.3.2 Information Structures

$\hat{h}_i^t$ is agent i's private history before t. This consists of an experimentation path $\{k_{i\tau} : \tau < t\}$ and news arrival times $\left\{T_{in}^{\theta,k} : T_{in}^{\theta,k} < t, n \in \mathbb{N}\right\}$ where $T_{in}^{\theta,k}$ is the arrival time[36] of agent i's n^{th} lump sum. $\hat{h}_i^0$ is a singleton. Agent i always knows their own private history.[37] Denote the profile of private histories by $h^t = \left(\hat{h}_1^t, \ldots, \hat{h}_N^t\right)$. Let $\hat{H}_i^t$ be the set of all private histories before t for agent i, and $H^t = \hat{H}_1^t \times \cdots \times \hat{H}_N^t$. The collection of all private histories for agent i is $\hat{H}_i = \cup_{t \geq 0} \hat{H}_i^t$, and the set of all profiles of private histories is $H = \cup_{t \geq 0} H^t$.

A (pure) *information structure* is a space of signals M and a profile of functions $\boldsymbol{\rho} = (\rho_i)_{i \in N}$ where

$$\rho_i : H \to M.$$

The designer chooses this information structure before time 0 and commits to it throughout the course of the game.[38] This choice is common knowledge among the agents. The interpretation of the information structure is that, for each possible history of the game, the designer sends agent i a signal from the set M according to ρ_i. For now, information structures are deterministic; each history maps uniquely into a profile of signals. While this is with loss of generality, this assumption will be relaxed in section 2.6.2.[39] Occasionally we shall, for linguistic and analogical convenience, call an information

[36] Formally, $T_{in}^{\theta,k} = \inf\left\{t : \nu_{it}^{\theta,k} \geq n\right\}$.

[37] Let $\left\{\hat{\mathcal{F}}_{it} : t \geq 0\right\}$ be the filtration generated by $\left\{\left(k_{it}, \nu_{it}^{\theta,k}\right) : t \geq 0\right\}$. Agent i knowing their own history formally means that their information at time t always contains $\hat{\mathcal{F}}_{it-}$.

[38] One can therefore think of the designer as observing all available information except for the state θ. However, commitment means that, for most of the game, she is non-strategic.

[39] Let Ψ be the space of all functions from H into M. A (mixed) information structure, or stochastic mechanism, is the space M along with a map $\phi : [0,1] \to \Psi^N$. The designer chooses (M, ϕ) and draws a uniformly distributed random variable whose realisation determines a (deterministic) profile $\boldsymbol{\rho}$. ϕ is therefore a Ψ^N-valued random variable. When we make use of stochastic mechanisms, we shall identify the mechanism with ϕ, and call profiles of functions $\boldsymbol{\rho}$, which are realisations of the

structure a *mechanism*. An information structure is *symmetric* when $\rho_i = \rho_j$ for all $i, j \in N$. Otherwise, the information structure is *private*.

Consider, for the sake of clarity, a pair of simple information structures. First, the designer could choose, if she wished, to reveal to agent i all information about the history of the game by choosing $M = H$, and letting ρ_i be the identity function.[40] Similarly, she could choose to reveal to agent i no information at all, other than i's own, by choosing M to be a singleton. Each choice of information structure induces a game of strategic experimentation among the agents. For example, if the information structure always revealed all information to all agents, the induced game—which shall be called the *perfect monitoring game*—is the classic setting of strategic experimentation analysed in Keller et al. (2005) and Keller and Rady (2010, 2015). An information structure that reveals no information at all induces a degenerate game where agents experiment in isolation as studied in Presman (1991).

Let $m_{it} \in M$ be the signal that i receives at time t. Agent i's *augmented history* h_i^t is a pair $\left(\hat{h}_i^t, m_i^t\right)$ where $m_i^t = \{m_{i\tau} : \tau < t\}$. H_i^t and H_i are defined as before. For simplicity, we will subsequently refer to an augmented history as just a history.

The space of mechanisms is vast, and the optimisation problems they induce for the agents are, in general, quite complicated. We thus restrict attention to a tractable class of mechanisms that includes many cases of interest[41] but also rules out many cases that are not[42]. In particular, we focus on Markov information structures in straightforward signals. We define these two terms presently.

Take an arbitrary history $h \in H$. The history h induces the designer's posterior belief p_D, which we identify with the probability that $\theta = 1$ conditional on h. We then say that an information structure is *Markov*[43] if it is a function of the designer's belief, and only the designer's belief. That is, for any pair of histories $h, h' \in H$ such that $p_D(h) = p_D(h')$, it must be the case that $\rho_i(h) = \rho_i(h')$ for all $i \in N$. In the sequel, we shall write ρ_i as a function of p_D alone.[44]

An information structure in *straightforward signals* is one where $M = [0, 1]$. A signal $\rho_i(h)$ is then taken to be an action recommendation for agent i after history h. The

random variable ϕ, disclosure rules. This definition is adopted from Aumann (1964). See Guo (2016) for the use of this definition in a mechanism design setting with Poisson bandits. Bonatti and Hörner (2017b) also make use of mixed policies in a model of strategic experimentation with public conclusive bad news but unobservable actions.

[40] The designer's signal to agent i under this information structure formally does not include the past signals sent to agents other than i. However, i can always compute what those signals were because they know the (deterministic) information structure.

[41] For example, any equilibrium constructed in Keller et al. (2005) and Keller and Rady (2010, 2015) can be induced in this class.

[42] As mentioned in the introduction, we wish to rule out the use of punishments. There does not appear to be a weaker assumption that suffices.

[43] Stationarity is implicit in this definition, as is common.

[44] We impose the technical restriction that each ρ_i is left- and piece-wise Lipschitz-continuous as functions of p_D. See Keller and Rady (2010) for details.

restriction to straightforward signals is standard from the literature on information design.[45] An information structure in straightforward signals shall be called a direct mechanism when convenient. We shall refer to any equilibrium (of the game induced by a direct mechanism) where $k_{it} = m_{it}$ for all $t \in [0, \infty)$ and $i \in N$ as *obedient*. If a mechanism in straightforward signals induces an obedient equilibrium, say that the mechanism is *implementable*.

2.3.3 Belief Dynamics: Designer

We first describe the dynamics of of the designer's belief. The results here are standard since the designer always observes all agents' histories. Agent belief dynamics depend on details of the information structure and the structure of news so their discussion is deferred to later sections.

Fix a profile of experimentation paths $\{k_{it} : t \geq 0, i \in N\}$. As long as no news has arrived, the designer's belief evolves deterministically according to

$$\dot{p}_D = -\left(\lambda_1 - \lambda_0\right) p_D \left(1 - p_D\right) \sum_{i \in N} k_i. \tag{2.3.1}$$

If news arrives at time t then beliefs jump[46] according to

$$p_{Dt+} = \frac{\lambda_1 p_{Dt}}{\lambda\left(p_{Dt}\right)}$$

where $\lambda(p) = p\lambda_1 + (1-p)\lambda_0$, and $x_{t+} = \lim_{t \downarrow 0} x_t$. Define

$$J(p) = \frac{\lambda_1 p}{\lambda(p)}$$

so that $J(p_{Dt})$ describes the jump in the designer's belief upon observing news at time t. In the conclusive case ($\lambda_0 = 0$), $J(p) = 1$ for any $p \in (0, 1)$.

2.3.4 Strategies and Equilibrium

Before describing the agents' strategies, let us first be explicit about the timeline of the game. Before the start of the basic game, nature chooses the state θ, and the designer

[45] However, as noted earlier, assuming this restriction here is stronger than is typical in the literature where it is often without loss as long as the designer can send random signals conditional on her information. As discussed in the introduction, the combination of this and the Markov assumption makes it difficult for the designer to discipline outcomes off the equilibrium path. As a consequence, there is still loss of generality *even if* the designer could send random signals. Unfortunately, it is undesirable to dispense with the Markov assumption, but the problem becomes intractable without restricting attention to straightforward signals. A general characterisation of implementable Markovian information structures would make an interesting future line of research.

[46] p_{Dt} is the belief before news arrives since we model beliefs to be left-continuous. p_{Dt+} is therefore the belief immediately after a jump.

chooses an information structure.[47] Then, at time t, each agent i receives a potentially private signal m_{it}, and then chooses an action k_{it} which determines the arrival rate of payoffs from the risky arm that they privately observe. Any information about other agents' actions and payoffs arrives only through the signal m_{it}.

A (pure) strategy $k_i : H_i \to [0,1]$ for agent i is a map from histories into actions.[48] Let $\{k_{it} : t \geq 0\}$ be the experimentation path induced by this strategy. Their expected continuation value at time t is given by

$$\mathbb{E}\left[\int_t^\infty re^{-r(\tau-t)}\left(h\,\mathrm{d}\nu_{i\tau}^{\theta,k} + (1-k_{i\tau})\,s\,\mathrm{d}\tau\right)\right].$$

A standard transformation, using the fact that compensated Poisson processes are martingales, and the Law of Iterated Expectations, allows us to write the above expression as

$$\mathbb{E}\left[\int_t^\infty re^{-r(\tau-t)}((1-k_{i\tau})s + k_{i\tau}\lambda\left(p_{D\tau}\right)h)\,\mathrm{d}\tau\right].$$

This expression is familiar from Keller and Rady (2010). The difference is that agent i does not directly observe p_{Dt} but only receives signals about it via m_{it}. However, the process $\{m_{it} : t \geq 0\}$, along with a conjecture about the other agents' strategies, induces beliefs about $\{p_{Dt} : t \geq 0\}$. Hence, a natural state variable for an agent's strategy is their beliefs over the designer's posterior. As a consequence, a strategy is now a map

$$k_i : \Delta\left([0,1]\right) \to [0,1]$$

from beliefs over the designer's posterior into actions. Such strategies shall be called *(stationary) Markov strategies.*

We thus aim to characterise obedient (weak perfect Bayesian) equilibria in Markov strategies induced by direct Markov mechanisms.[49] These will be referred to as simply equilibria. Note, in particular, that when we speak of equilibria, they are obedient, unless specified otherwise.

2.4 Conclusive Good News

We first consider an environment where news is good and conclusive so $\lambda_1 h > s > \lambda_0 h = 0$. The mechanism that reveals all information in this setting is studied in Keller et al. (2005) whose results we make use of below. In particular, they derive belief cut-offs p_N^*; $p_N^\dagger$; p_1^*; p^m; and, the value function $W^\dagger(p)$. p_N^* is the solution to the N-agent cooperative problem that maximises all players' average utility. All agents

[47] The order of these choices does not matter, since nature is non-strategic, and the designer does not observe θ.

[48] To save on notation, we use k_i to refer to both i's strategy, and their action at some given time.

[49] The restriction to obedient equilibria is without loss, by exactly the argument in Forges (1986), proposition 1.

choose $k = 1$ above this cut-off and $k = 0$ below it. p_1^* is therefore the solution to the single-agent problem while $p_N^\dagger$ is the cut-off characterising the symmetric Markov-perfect equilibrium[50] (MPE) under perfect monitoring whose value function is given by $W^\dagger(p)$. In this MPE, all agents choose $k = 1$ above $p_N^\dagger$, and choose some interior $k \in (0,1)$ below it until they reach the belief p_1^* at which point all experimentation stops. p^m is the myopic cut-off belief, below which a myopic agent never experiments since it entails negative flow payoffs. We have that $p_N^* < p_1^* < p_N^\dagger < p^m$. The details of these results are summarised in appendix 2.8.

2.4.1 Belief Dynamics: Agent

We wish to simplify an agent's problem in order to apply standard dynamic programming techniques. To this end, we state a sequence of lemmas that characterise agents' equilibrium belief dynamics. Recall that p_1^* is the belief cut-off that an agent uses in autarky, so that if the designer gives agent i no information, their unique optimal strategy is to choose $k_i = 1$ at all beliefs above p_1^* and $k_i = 0$ at all beliefs below p_1^*.

Lemma 2.1. *A direct Markov mechanism induces an obedient equilibrium only if exactly one of the following two conditions hold for each $i \in N$:*

1. *$\rho_i(1) = 1$; or,*
2. *ρ_i is identically zero and $p_0 \leq p_1^*$.*

Proof. Consider a private history for agent i where they observe a success. Conditional on this history, $p_D = p_i = 1$. The uniquely optimal action for i is $k_i = 1$. Thus, if $\rho_i(1) < 1$, there exists some history where $\rho_i \neq k_i$. This history must occur with positive probability, or else the designer never tells agent i to experiment. In the former, choosing $\rho_i(1) < 1$ contradicts obedience. In the latter, when the designer never tells agent i to experiment, so it is without loss to assume that $\rho_i = 0$, the agent is obedient if and only if $p_0 \leq p_1^*$. □

This result is a consequence of the Markovian structure of the designer's signals, which limits the ways she can communicate the history of the basic game to each agent. In particular, guaranteeing obedience from agents who have observed successes forces her to recommend experimentation with full intensity to all agents. This does not, however, enforce symmetry of signals at all beliefs. In general, the designer is still free to recommend different intensities of experimentation, or, indeed, no experimentation at all, to different agents for interior beliefs.[51]

[50] With perfect monitoring, a Markov-perfect equilibrium is a subgame-perfect equilibrium in Markov strategies that use the common posterior as the state variable.

[51] Note that lemma 2.1 does not make use of the fact that we have assumed that mechanisms are deterministic. The reasoning of the proof applies to every disclosure rule that the designer employs with positive probability under a stochastic mechanism. We shall make use of this observation in section 2.6.2.

The next lemma gives a partial characterisation of an agent's belief over the designer's posterior.

Lemma 2.2. *In any equilibrium, an agent's belief over the designer's posterior has at most two points in its support:* 1, *and some* $q_t \in (0, p_0)$ *which evolves according to*

$$\dot{q} = -\lambda_1 q(1-q)\left[k_i + \sum_{j\neq i} \rho_j(q)\right]. \tag{2.4.1}$$

Proof. If any agent has observed a success, then $p_D = 1$. However, if no agent has succeeded, then the designer's beliefs evolve deterministically, and also determine deterministic paths $\{k_{it} : t \geq 0, i \in N\}$ that, by obedience, induce the dynamics given above. These two events are mutually exclusive and exhaustive so must therefore fully specify the designer's possible beliefs. □

The designer's posterior always drifts downward in a deterministic fashion in the absence of a success. This allows agents, who are not certain about the state, to compute, as shown in the lemma, this posterior in the event that none of agents has enjoyed success.[52] We shall call this object the designer's *worst-case belief*, and continue to denote it by q. Note that q represents what the agents believe the designer's posterior is conditional on her not having observed a success; it is not the designer's actual posterior. Nevertheless, conditional on the absence of news, the designer's posterior will indeed be q in equilibrium.

A version of lemma 2.2 is also true even when the equilibrium is not obedient. We only need to replace $\sum_{j\neq i} \rho_j(q)$ with the agent's (equilibrium) conjecture $\sum_{j\neq i} k_i$ when $p_D = q$. The proof of this result is similar.

Corollary 2.1. *In equilibrium,* $\rho_i \in [0, 1)$ *implies that* $p_i = p_D$*, unless* ρ_i *is identically zero.*

Proof. Lemma 2.1 establishes that when $\rho_i \neq 1$ and is not identically zero, $p_D < 1$. Lemma 2.2 then shows that agent i's belief (over the designer's posterior) is degenerate. Equilibrium guarantees that this belief is correct. □

This corollary says that any recommendation other than full experimentation perfectly reveals the designer's beliefs, and so any agent who receives such recommendations must have a posterior that coincides with the designer's. That is, recommending partial

[52]This argument clearly makes strong use of our assumption that the designer and the agents all use pure strategies, which obviates strategic uncertainty. While unsatisfactory, allowing for general mixed strategies introduces significant difficulty in the analysis, as demonstrated by Bonatti and Hörner (2017b). See, however, 2.6.2.

(or no) experimentation discloses the designer's (payoff-relevant) information. This is because any recommendation other than full experimentation can only be consistent with the absence of news. However, since the agents know what the designer's belief is conditional on that event, they must therefore know this belief once the said event has been revealed to be the case.

One might be tempted to think that an agent can always just invert their ρ_i in order to discover the designer's belief, therefore delivering a stronger version of the above result. However, we have not assumed enough structure on the ρ_i's to guarantee that these inverses are well-defined, which means that this method is not guaranteed to uniquely pin down the designer's belief.[53] Nevertheless, the results above guarantee, as we shall see, that the designer can only exploit the non-invertibility of the ρ_i's, if at all, in a limited way.

The next result derives the dynamics of an agent's belief whenever their posterior does not coincide with the designer's.

Lemma 2.3. *On any interval of time $[t, t + \mathrm{d}t)$ on which $\rho_i = 1$, i's belief p_i, conditional on not observing a success, evolves according to*

$$\dot{p}_i = -\lambda_1 p_i (1 - p_i) k_i. \tag{2.4.2}$$

Proof. This comes from an application of Bayes' rule. Note that on any interval $[t, t + \mathrm{d}t)$ with $\rho_i = 1$, the designer's signal provides no information about θ. Hence, i's posterior depends only on their own experimentation. More explicitly:

$$p_{i,t+\mathrm{d}t} = \left[p_{it}e^{-\lambda_1 \sum_{j\neq i} k_{jt}\mathrm{d}t} + 1 - p_{it}\right] \frac{p_{it}e^{-\lambda_1 \sum_{j\in N} k_{jt}\mathrm{d}t}}{p_{it}e^{-\lambda_1 \sum_{j\in N} k_{jt}\mathrm{d}t} + (1 - p_{it})} + p_{it}\left(1 - e^{-\lambda_1 \sum_{j\neq i} k_{jt}\mathrm{d}t}\right).$$

Subtracting p_{it} from both sides, dividing by $\mathrm{d}t$, and taking $\mathrm{d}t \to 0$ gives the result. □

To understand the equation in the proof above, notice that the right hand side gives the i's expected posterior at the end of the interval $[t, t + \mathrm{d}t)$. The expression in brackets in the first term is the probability that no other agent observes a success, which is multiplied by the agent's posterior conditional on this event. The second term is just the probability that some other agent does observe a success, in which case i's posterior jumps up to 1. Consequently, whenever the designer tells i to experiment fully, their only source of information about θ is their own experimentation.

2.4.2 Equilibrium Characterisation

The previous results and the associated discussion give us the following theorem.

[53] However, piece-wise Lipschitz-continuity means that, by the implicit function theorem, each ρ_i is locally invertible almost everywhere, which gives us another way of viewing the above results.

Theorem 2.1. *With conclusive good news, any obedient equilibrium where all agents experiment with positive probability is characterised by at most 2^N phases. Each agent must be in one of two mutually exclusive phases: one where the designer discloses no information and recommends full experimentation, and another where the designer reveals all (payoff-relevant) information and recommends less than full experimentation.*

The above result does not require all phase transitions to be permanent. The designer can, in principle, after a period of recommending less than full experimentation, switch to recommending full experimentation and vice-versa. However, once the designer instructs all agents to stop using the risky arm, all experimentation ceases, and the game (effectively) ends. This is because the designer's belief becomes frozen[54] at any point where she only recommends the safe arm. Notice that this instance is just a special case of the phase where the designer completely discloses all information, and is thus implicit in the above characterisation.[55]

Whenever the agents are told to experiment fully, then their beliefs decay more slowly than the designer's worst-case belief.[56] This enables the designer to induce the agents to experiment more than she would have in the perfect monitoring case.[57] Nevertheless, there is an additional force at work: agents wish to continue experimenting fully because it speeds up the arrival of information. Disobeying the designer simply slows down the rate at which her worst-case belief decays, which delays the transition into the alternative phase of full information revelation.

Given the agents' belief dynamics, we can now write the Hamilton-Jacobi-Bellman equations of the agents' maximisation problems.

2.4.3 Optimality Equations

From theorem 2.1 we know that, in an obedient equilibrium, agent i either knows exactly the designer's belief, or is uncertain about whether the designer has observed a success.[58] It follows then that agent i's value function must be characterised by a pair of Hamilton-Jacobi-Bellman equations depending on which regime they are in.

[54] See (2.3.1).

[55] Another boundary case worth mentioning is when there is only a single agent experimenting fully for an open interval of time. In this case, while the agent knows that their posterior decays at the same rate as the designer's (if hers does at all), these posteriors may not coincide. Moreover, the designer reveals no information because there is no information being generated to reveal.

[56] Compare (2.4.1) and (2.4.2).

[57] Heidhues et al. (2015) exploit a similar observation in a discrete-time version of this game, where they make payoffs private, but allow for communication. They show that socially efficient experimentation can be achieved as a sequential equilibrium for some parameterisations of the aforementioned game. However, the equilibria they construct rely on punishments to deter deviations, which we have ruled out.

[58] We ignore the trivial case where the designer recommends no experimentation at all.

When agent i's belief coincides with the designer's, we have the familiar optimality equation[59]

$$rV_i = rs + \max_{k\in[0,1]} k\left\{ r\left[p\lambda_1 h - s\right] + \lambda_1 p\left[\lambda_1 h - V_i - (1-p)\partial_p V_i\right]\right\} \qquad (2.4.3)$$
$$+ \lambda_1 p\left[\lambda_1 h - V_i - (1-p)\partial_p V_i\right]\sum_{j\neq i}\rho_j\left(p\right).$$

This is exactly the HJB equation that appears in Keller et al. (2005). As they describe, experimentation gives a flow payoff, captured by $r\left[(1-k)\,s + kp\lambda_1 h\right]$. The term $k\lambda_1 p\left[\lambda_1 h - V_i - (1-p)\partial_p V_i\right]$ captures i's expected information gain from their own experimentation, which consists of two parts. First, at rate $k\lambda_1 p$, a success arrives, which yields a gain of $\lambda_1 h - V_i$. Otherwise, i's value decays according to $-k\lambda_1 p(1-p)\partial_p V_i$ due to a decline in their belief. The remaining term describes i's benefit from other agents' experimentation.

When agent i's beliefs do not coincide with the designer's, their value is a function of their beliefs over the designer's posterior. By the law of iterated expectations, $p_i = \mathbb{E}\left[p_D\right]$. Thus, what agent i thinks about the designer's belief can be characterised by two parameters, p_i, their own posterior about the state of the world, and q, the designer's worst-case belief. To see this, suppose that i has not observed a success by time t, and let E_t denote the event that no one else has observed a success either. Conditional on E_t, the designer's posterior must be given by q_t. Since the agent's information is always subsumed in the designer's, we must have that

$$p_{it} = \Pr\left(E_t\right) q_t + \left[1 - \Pr\left(E_t\right)\right]. \qquad (2.4.4)$$

Thus, the agent's beliefs over the designer's posterior is uniquely determined by p_{it} and q_t. We use these as our state variables.

We therefore obtain a similar but slightly less familiar equation

$$rV_i = rs + \max_{k\in[0,1]} k\left\{ r\left[p\lambda_1 h - s\right] + \lambda_1 p\left[\lambda_1 h - V_i - (1-p)\partial_p V_i\right]\right. \qquad (2.4.5)$$
$$\left. - \lambda_1 q\left(1-q\right)\partial_q V_i\right\} - \lambda_1 q\left(1-q\right)\partial_q V_i \sum_{j\neq i}\rho_j\left(q\right).$$

This equation features the additional terms $-k\lambda_1 q\left(1-q\right)\partial_q V_i$ and $-\lambda_1 q\left(1-q\right)\partial_q V_i \sum_{j\neq i}\rho_j\left(q\right)$, which now capture the benefit of bringing forward the information disclosure phase induced by the decay of the designer's worst-case belief. As before, the former expression captures the benefit from i's own experimentation, while the latter captures the benefit

[59] If the derivatives shown above are interpreted in a generalised sense (Clarke, 1990), then the Hamilton-Jacobi-Bellman equations are both necessary and sufficient. However, it turns out that these equations have solutions even when the derivatives are given their classical definitions, so considering generalised derivatives is not necessary. See Davis (1993) for details.

from other agents' experimentation. However, i's payoff from other agents' actions does not include a jump term proportional to $\lambda_1 h - V_i$, as it did in the (2.4.3), where i knew the designer's posterior, since i does not observe other agents' successes.

2.4.4 Symmetric Monotone Mechanisms

In order to get a sharper characterisation of equilibria, we assume that the mechanism is symmetric, so that $\rho_i = \rho_j = \rho$ for all $i, j \in N$, and monotone. Monotonicity means that ρ is non-decreasing in the designer's belief.[60] The following result is immediate from theorem 2.1.[61]

Corollary 2.2. *When news is good and conclusive, any equilibrium induced by a symmetric monotone mechanism is characterised by a cut-off belief $\tilde{q}$. Above this cut-off, the designer discloses no information and recommends full experimentation. At and below the cut-off, the designer reveals all information and recommends the unique symmetric (Markov-perfect) equilibrium action of the perfect monitoring game.*

Thus, in the symmetric monotone case, mechanisms that induce obedient equilibria admit a simple characterisation. First, the designer recommends full experimentation to all agents. Once the designer's posterior has reached the cut-off identifying the mechanism, she must reveal her information. To be specific, at the time agents expect the designer's posterior to cross $\tilde{q}$, the designer's next recommendation reveals whether a success has arrived in the past or not. If she continues to recommend full experimentation, agents conclude that a success must have arrived. Otherwise, they conclude that her true posterior must be given by $\tilde{q}$. Keller et al. (2005) then establish that there is a unique symmetric MPE in this setting, so that MPE determines the designer's action recommendation.

We now proceed to a characterisation of the cut-off $\tilde{q}$ that implements the equilibria described above. We begin with an easy observation.[62]

Corollary 2.3. *A cut-off $\tilde{q}$ induces an obedient equilibrium only if $\tilde{q} \leq p_N^\dagger$.*

Proof. Suppose that $\tilde{q} > p_N^\dagger$, so that the designer stops recommending complete experimentation at a belief strictly above $p_N^\dagger$. At these beliefs, the designer's posterior is now common knowledge. However, by proposition 2.3, $k = 1$ is the unique best response at these beliefs. This contradicts obedience. □

[60] If ρ were non-increasing, then lemma 2.1 implies that ρ must be identically 1 or identically 0. The former cannot induce an obedient equilibrium, while the latter induces only trivial equilibria with $p_0 \leq p_1^*$.

[61] See the beginning of this section for a brief summary of the perfect-monitoring MPE. Relevant details are in appendix 2.8.

[62] In what follows, we assume that $\tilde{q} \leq p_0$. This is clearly without loss, since choosing $\tilde{q} > p_0$ is the same as choosing $\tilde{q} = p_0$.

2.4.4.1 Solution to HJB (2.4.9)

Consider the phase where the designer reveals no information and recommends the action $k=1$, and suppose all agents other than i obey this recommendation. When agent i follows this recommendation as well, their value function satisfies

$$rV_i = p\lambda_1 \left[rh + \lambda_1 h - V_i - (1-p)\,\partial_p V_i\right] - N\lambda_1 q\,(1-q)\,\partial_q V_i.$$

This is a first-order linear partial differential equation, which can be solved using the method of characteristics.[63] The solution is given by

$$V_i\,(p,q) = p\lambda_1 h + (1-p)\,\Omega\,(p)^{\frac{r}{\lambda_1}}\,F\left(\frac{\Omega\,(p)^N}{\Omega\,(q)}\right), \tag{2.4.6}$$

where $\Omega\,(p) = \frac{1-p}{p}$ and F is a C^1 function. In order to determine F, we need an appropriate boundary condition, as given by

$$V_i\,(p,\tilde{q}) = \frac{1-p}{1-\tilde{q}}W^{\dagger}\,(\tilde{q}) + \frac{p-\tilde{q}}{1-\tilde{q}}\lambda_1 h.$$

To see this, observe that, from (2.4.4), we have[64]

$$\Pr\,(E_t) = \frac{1-p_t}{1-q_t}. \tag{2.4.7}$$

Suppose $q_t = \tilde{q}$. With the probability $\Pr\,(E_t)$, no success has arrived, so the designer stops recommending full experimentation, and instructs the agents to play the symmetric perfect-monitoring MPE instead. The agents' continuation value then must be $W^{\dagger}\,(\tilde{q})$, the common value function in the said MPE. With complementary probability, the designer must have observed a success. The agents are able to infer this from the fact that the designer continues to recommend $k=1$, instead of the MPE action. Thus their continuation value must be $\lambda_1 h$.

Using this boundary condition we find that

$$F\,(x) = \left[\frac{W^{\dagger}\,(\tilde{q}) - \tilde{q}\lambda_1 h}{(1-\tilde{q})\,\Omega\,(\tilde{q})^{\frac{r}{N\lambda_1}}}\right] x^{-\frac{r}{N\lambda_1}}.$$

Hence, i's value function from obeying the designer's recommendations is

$$V_i\,(p,q) = p\lambda_1 h + (1-p)\,\frac{W^{\dagger}\,(\tilde{q}) - \tilde{q}\lambda_1 h}{1-\tilde{q}}\left[\frac{\Omega\,(q)}{\Omega\,(\tilde{q})}\right]^{\frac{r}{N\lambda_1}}. \tag{2.4.8}$$

This value function is the sum of payoffs from the risky arm, plus the benefit of inform-

[63] See Evans (2010).

[64] See section 2.9.2 for a method of deriving the expression below without appealing to the law of iterated expectations.

ation from experimentation. Notice that when $p = q = p_0$, and $\tilde{q} = p_N^\dagger < p_0$, the value function above is exactly equal to an agent's ex-ante payoff in the perfect monitoring MPE, as given in proposition 2.3. On the other hand, when $\tilde{q} = p_N^*$, the above value is almost identical to the value of the N-agent cooperative problem. Indeed, when $p = q = p_0$, this gives the agents' ex-ante payoff under the efficient solution. Moreover, we have that

$$-N\lambda_1 q\,(1-q)\,\partial_q V_i = r\,(1-p)\,\frac{W^\dagger(\tilde{q}) - \tilde{q}\lambda_1 h}{1-\tilde{q}}\left[\frac{\Omega(q)}{\Omega(\tilde{q})}\right]^{\frac{r}{N\lambda_1}},$$

which we shall discover is positive[65], confirming the benefit of inducing decreases in the worst-case belief.

2.4.4.2 Verifying Obedience

Agent i has no incentive to deviate if and only if, for each pair (p, q) along the equilibrium path, $k_i = 1$ solves the maximisation problem in (2.4.5), where V_i is given in (2.4.8), and $\rho_j = 1$ for $j \neq i$. In other words, the expression in braces in (2.4.5) must be greater than or equal to zero. A straightforward calculation shows that this is equivalent to

$$p\lambda_1 h + \frac{1-p}{N}\frac{W^\dagger(\tilde{q}) - \tilde{q}\lambda_1 h}{1-\tilde{q}}\left[\frac{\Omega(q)}{\Omega(\tilde{q})}\right]^{\frac{r}{N\lambda_1}} \geq s. \tag{2.4.9}$$

The interpretation of the condition above is standard. The left-hand side is the marginal benefit to experimentation, which consists of the flow payoff $p\lambda_1 h$ and the benefit from a decline of the worst-case belief, as given in the second term. The latter forms $\frac{1}{N}$ of the total value of information found in the second term of (2.4.8), which is exactly i's contribution to this benefit. This can be thought of as a kind of *encouragement effect*, as mentioned in the introduction.[66] Agents are encouraged to experiment because it brings forward the date of information disclosure from the designer. It is exactly this effect that enables the designer to induce experimentation for longer than agents would have in the perfect monitoring case. The right hand side is the marginal cost of experimentation, which is just the opportunity cost of not using the safe arm.

The inequality (2.4.9) must hold for each p and q along the equilibrium path, and therefore determines a continuum of constraints to check. This exercise is simplified by the fact that, along this path, p and q are related by the equality[67]

$$\frac{\Omega(q)}{\Omega(p_0)} = \left[\frac{\Omega(p)}{\Omega(p_0)}\right]^N, \tag{2.4.10}$$

along with the observation below, which is straightforward to verify.

[65] This is shown in the proof of proposition 2.1 in appendix 2.9.3.
[66] The properties asserted there are now easy to verify.
[67] See appendix 2.9.2 for a derivation.

Proposition 2.1.

$$MB(p;\tilde{q}) = p\lambda_1 h + \frac{1-p}{N}\frac{W^{\dagger}(\tilde{q}) - \tilde{q}\lambda_1 h}{1-\tilde{q}}\left[\frac{\Omega(p_0)}{\Omega(\tilde{q})}\right]^{\frac{r}{N\lambda_1}}\left[\frac{\Omega(p)}{\Omega(p_0)}\right]^{\frac{r}{\lambda_1}} \tag{2.4.11}$$

is strictly convex in p*. Moreover, it is strictly decreasing in* p *within a neighbourhood of* 0*, and strictly increasing in* p *within a neighbourhood of* 1*. Finally, as a function of* $\tilde{q}$*,* $MB(p;\tilde{q})$ *is strictly increasing on* $(0, p_N^*)$ *and strictly decreasing on* $\left(p_N^*, p_N^{\dagger}\right)$*.*

The convexity of $MB(p;\tilde{q})$ is inherited from the convexity of the value function V_i upon substituting for $\Omega(q)$ as in (2.4.10). This convexity reflects the value of information: agents benefit from mean-preserving gambles over their posterior. On the other hand, that $MB(p;\tilde{q})$ is increasing in $\tilde{q}$ in the direction of p_N^* comes reflects the fact that experimentation has the greater value the closer the cut-off is to the socially efficient one. In other words, the encouragement effect is largest at efficiency.

Proposition 2.1 is useful because it reduces the problem of checking a continuum of constraints into verifying pairs of (in)equalities which constitute necessary and sufficient conditions for the implementability of $\tilde{q}$. These conditions are stated in appendix 2.9.4. We use one of these sufficient conditions to derive the next result.

Theorem 2.2. *The cut-off* $\tilde{q} = p_N^{\dagger}$ *is implementable. When* $p_0 > p_N^{\dagger}$*, the set of implementable cut-offs contains an interval* $\left[\hat{p}, p_N^{\dagger}\right]$*, where* $\hat{p} < p_N^{\dagger}$*.*

Of course, we already knew from the revelation principle that the cut-off $p_N^{\dagger}$ must be implementable, since the designer's recommendations when $\tilde{q} = p_N^{\dagger}$ are exactly what the agents would have done in the symmetric MPE of the perfect monitoring game. However, it is reassuring that this fact can be proven directly. Moreover, the direct proof gives us the second part of the theorem, which cannot be deduced from the revelation principle alone.

Proposition 2.1 also delivers the next corollary. $MB(p;\tilde{q})$ is increasing in $\tilde{q}$ on $(0, p_N^*)$, so implementing a higher cut-off in this range assuming a lower cut-off were implementable poses no problem for satisfying (2.4.9).

Corollary 2.4. *If* $\tilde{q} < p_N^*$ *is implementable, so is every* $q' \in [\tilde{q}, p_N^*]$*.*

The corollary can also be strengthened in the following way.

Theorem 2.3. *Suppose* $\tilde{q} < p_N^{\dagger}$ *is implementable. There exists an* $r_0 > 0$ *such that for all* $r \in (0, r_0)$*, every* $q' \in \left[\tilde{q}, p_N^{\dagger}\right]$ *is also implementable. Hence, for all* r *small enough, the set of implementable cut-offs forms an interval contained in* $\left[0, p_N^{\dagger}\right]$*.*

The condition on r is used to guarantee that MB is 'monotonic enough' as cut-off $\tilde{q}$ increases. Without it, increasing $\tilde{q}$ on $\left(p_N^*, p_N^{\dagger}\right)$ could lead to a violation of (2.4.9), since MB is strictly decreasing on that interval. Numerical calculations suggests that r need not be too small for the above result to be true.

2.4.4.3 Transparency and Efficiency

The above discussion identifies a partial trade-off between transparency and efficiency. Observe that, by proposition 2.1, V_i is strictly decreasing in $\tilde{q}$ on $\left(p_N^*, p_N^\dagger\right)$, and strictly increasing on $\left(0, p_N^*\right)$.[68] Since the designer reveals all information below $\tilde{q}$, and conceals information above, one can view the choice of a higher $\tilde{q}$ as choosing a more transparent mechanism, and the choice of a lower $\tilde{q}$ as, naturally, choosing a less transparent one. As a consequence, on $\left(p_N^*, p_N^\dagger\right)$, the choice between transparency and efficiency is zero-sum; choosing to increase one must, by necessity, decrease the other. On the other hand, on $\left(0, p_N^*\right)$, there is no trade-off. Transparency and efficiency go hand-in-hand.

It is natural to ask, then: when can the designer achieve efficiency? When is she restricted to complete transparency, as represented by $\tilde{q} = p_N^\dagger$? We turn to these questions next. We begin with a simple answer to the latter.

Theorem 2.4. *For every $p_0 \in (0, p^m)$, where $p^m = \frac{s}{\lambda_1 h}$, there exists an $N_0 > 1$ such that for all $N \geq N_0$, the uniquely implementable cut-off is $p_N^\dagger$.*[69]

Proof. This result is obtained by inspecting (2.4.9). When $p_0 < p^m$, we must have $s - p\lambda_1 h > 0$ for any possible on-path belief. However, the second term on the left-hand side of (2.4.9) can be made as small as possible by choosing N to be large enough, from which we obtain a violation of the inequality (2.4.9). □

This means that when experimentation is costly, in the sense that it entails negative flow costs[70], and there are a large enough number of agents, then the only implementable symmetric mechanism is one that reveals all information. The economics behind this result becomes transparent from examining $MB\left(p; \tilde{q}\right)$. When N is large, i's contribution to the decay of the worst-case belief q, which drives our encouragement effect, is small, and therefore cannot outweigh the cost of experimentation $s - p\lambda_1 h$. Hence, i is tempted to free-ride on other agents' experimentation, and the designer cannot induce them to use the risky arm.

Next we state a sufficient condition for implementing any cut-off that lies in $\left[p_N^*, p_1^*\right]$. This condition is just an application of proposition 2.5 in appendix 2.9.4. Recall that in the symmetric perfect-monitoring MPE, all experimentation stops once the common public belief reaches p_1^*. Hence, cut-offs in this region always induce more efficient

[68] The derivatives of $MB\left(p; \tilde{q}\right)$ and V_i with respect to $\tilde{q}$ are constant multiples of each other, so the same comparative static results apply.

[69] This does not contradict theorem 2.2, since $p_N^\dagger \to p^m$ as $N \to \infty$. To see this, observe that $p_N^\dagger$ is strictly increasing in N but must be bounded above by p^m.

[70] To understand why the assumption above about the prior belief means experimentation is costly, notice that the expected flow payoff from choosing $k > 0$ is $(1-k)\,s + kp\lambda_1 h$, which is negative for $p < p^m$. Thus, for these beliefs, the only reason to experiment would be the information value it provides, which is small for large enough N.

levels[71] of experimentation, in both amount and intensity.[72] We denote by $\hat{p}(q)$ the agents' equilibrium belief about the state when the designer's worst-case belief is at q. That is, for any q, $p = \hat{p}(q)$ satisfies (2.4.10).

Theorem 2.5. *The cut-off* $\tilde{q} \in [p_N^*, p_1^*]$ *is implementable if the following chain of inequalities hold:*

$$\lambda_1 h \left[\frac{\lambda_1 \hat{p}(\tilde{q})}{\lambda_1 \hat{p}(\tilde{q}) + r}\right] \geq \frac{1}{N}\left(\frac{s - \tilde{q}\lambda_1 h}{1 - \tilde{q}}\right) \geq \frac{s - \hat{p}(\tilde{q})\lambda_1 h}{1 - \hat{p}(\tilde{q})}.$$

The last inequality is necessary.

A simple way to guarantee all three inequalities is for the prior belief p_0 to be large enough. This is because $\hat{p}$ is increasing in p_0, and so increases the leftmost expression but decreases the rightmost expression.[73]

2.4.5 Relaxing Symmetry and Monotonicity

We end this section with a brief note on relaxing symmetry and monotonicity. One simple way to relax both assumptions simultaneously is in the choice of continuation equilibrium after the designer's belief has reached the threshold $\tilde{q}$. Below this threshold, let the designer recommend an asymmetric MPE, such as those constructed in Keller et al. (2005). One can also straightforwardly construct asymmetric but monotone mechanisms of the type where the designer's recommendations are extremal: $\rho_i \in \{0, 1\}$ for all $i \in N$. This class of mechanisms is now characterised by a collection $\{\tilde{q}_i : i \in N\}$ of cut-offs, one for each agent. $\tilde{q}_i$ is then the belief at which the designer tells agent i to switch to the safe arm, at least until news arrives. This construction amplifies the encouragement effect for agents who continue experimenting later than others, since the effect will now only be divided by the number of agents still experimenting at the worst-case belief.

Consider instead non-monotone, but symmetric, mechanisms. Given theorem 2.1, according to the same reasoning as in corollary 2.2, this class of mechanisms is characterised by a finite[74] collection of cut-offs $\{\tilde{q}_\ell\}$, where each cut-off determines a transition from full disclosure to full concealment and vice-versa. Our results above in the monotone case then characterise the lowest of these cut-offs. The strategies immediately to

[71] Keller et al. (2005) show that the amount of experimentation can be identified with the posterior at which experimentation stops. Intensity is a measure of how long it took to reach that posterior.

[72] In fact, Heidhues et al. (2015) show that, in a discrete-time version of the perfect monitoring game, every sequential equilibrium features no experimentation below p_1^*, so these cut-offs induce outcomes that are more efficient than any outcome under perfect monitoring that can be approximated by equilibria of the discretised game.

[73] This is similar to the result by Heidhues et al. (2015), but, as mentioned previously, they use punishments to support these outcomes.

[74] This is from our restriction that ρ must be piecewise Lipschitz-continuous.

the right of the lowest cut-off can then be determined in a manner similar to the construction of Keller et al. (2005).[75] One can then proceed as above to study the second lowest cut-off, and so on.

2.5 Conclusive Bad News

We consider now a model where news is bad and conclusive: $0 = \lambda_0 h > s > \lambda_1 h$. One might reasonably expect that the results in this case are just the mirror image of the results derived above. This is, to some extent, true, but not in all respects, as Keller and Rady (2015) show in the perfect monitoring case. We will find something similar here.

First we recall some results from the perfect information setting. As in the conclusive good news case, p_1^* and p_N^* characterise the single-agent and cooperative solution respectively. Now, however, the risky arm is used exclusively at all beliefs *below* the relevant cut-offs, and the safe arm is used above them. Of course, we must have that $p_1^* < p_N^*$. The conclusive bad news case also has a unique symmetric MPE, characterised by two cutoffs, $p_N^\dagger$ and $\bar{p}_N$, with $p_N^\dagger < p_1^* < \bar{p}_N < p_N^*$. Below $p_N^\dagger$, the risky arm is used exclusively. Between $p_N^\dagger$ and $\bar{p}_N$, agents allocate resources to both arms. Finally, above $\bar{p}_N$, only the safe arm is used.

2.5.1 Belief Dynamics: Agent

Now we state analogues of the results of from the previous section.

Lemma 2.4. *A direct Markov mechanism induces an obedient equilibrium only if exactly one of the following two conditions hold for each $i \in N$:*

1. *$\rho_i(1) = 0$; or,*
2. *ρ_i is identically zero and $p_0 > p_1^*$.*

When news is bad, the designer recommends that agents stop experimenting when she is sure that the risky arm is bad. The reason for this is the same as with good news. In order to satisfy obedience for agents who are certain of the state of the world, the designer must recommend the safe arm to agents who are also still uncertain. However, this means that the only agents the designer can conceal the arrival of news from are those agents who are not experimenting at all at the time of arrival of news. All other agents are immediately able to infer the arrival of bad news when the designer instructs them to switch to the safe arm. Of course, the agents who remain uninformed because

[75] Of course, the boundary conditions must be determined appropriately, as we did in the monotone case.

they were using the safe arm only at the arrival time of news eventually learn the state as well, as long as the designer instructs them to experiment at some (optimistic enough) beliefs.

Lemma 2.5. *In any equilibrium of the game with conclusive bad news, $\rho_i \in (0,1]$ implies that $p_i = p_D$, unless ρ_i is identically zero. When $\rho_i = 0$, p_i remains stationary, and agents' beliefs over the designer's posterior is supported on 1 and q, where q evolves according to*

$$\dot{q} = -\lambda_1 q(1-q) \sum_{j \neq i} \rho_j\left(q\right).$$

As in the case with conclusive good news, there is a full disclosure phase, and a phase where the designer reveals no information. The difference is that with conclusive bad news, the disclosure phase occurs any time the designer recommends any amount of experimentation. Agents learn nothing about the state otherwise, because they are instructed to use the safe arm.

2.5.2 Equilibrium Characterisation

We can use similar arguments to those from section 2.4 to obtain the following result.

Theorem 2.6. *With conclusive bad news, any obedient equilibrium is characterised by at most $2^N + 1$ phases. Except for the final phase, each agent must be in one of two mutually exclusive phases: one involving no information disclosure along with no experimentation; and, another where the designer reveals all (payoff-relevant) information while simultaneously recommending a strictly positive amount of experimentation. The final phase is when all experimentation stops.*

Let us now focus on symmetric monotone mechanisms. However, with bad news, a monotone mechanism is *non-increasing* in the designer's posterior. This gives us a result which starkly distinguishes the good news setting from that with bad news.

Theorem 2.7. *When $p_0 \leq p_1^*$, or $p_0 \geq \bar{p}_N$, there is a unique implementable symmetric monotone mechanism, identified by the unique symmetric MPE of the perfect monitoring game. When $p_0 \in \left(p_1^*, \bar{p}_N\right)$, there are two implementable symmetric monotone mechanisms, one that implements the symmetric MPE, and another that recommends no experimentation at all.*

Proof. Fix any $p_0 \in (0,1)$, and suppose that $\rho\left(p_0\right) > 0$. Monotonicity then requires that $\rho\left(p_D\right) > 0$ for any $p_D \in [0, p_0]$. The designer's belief dynamics (2.3.1) and theorem 2.6 then guarantees that the agents must be playing a game of perfect monitoring. Symmetry and obedience implies that they must be playing the symmetric MPE, which pins down the function ρ, as long as $p_0 < \bar{p}_N$. When $p_0 \geq \bar{p}_N$, the assumption $\rho\left(p_0\right) > 0$

contradicts obedience, since the unique best response in the perfect monitoring game at the belief p_0 is to choose $k = 0$. Hence, $\rho(p_0) = 0$, as in the symmetric perfect monitoring MPE.

Suppose $p_0 \leq p_1^*$, and that $\rho(p_0) = 0$. Every agent has a profitable deviation, since it is individually optimal to use the risky arm at p_0. When $p_0 \in (p_1^*, \bar{p}_N)$, recommending no experimentation at all is also implementable, since it is individually optimal and the designer never reveals any information in this equilibrium. □

Unlike in the case of good news, the set of implementable symmetric monotone mechanisms, when identified with the equilibria they induce, is not much larger then the set of symmetric perfect monitoring MPE. Furthermore, in the good news case, this set of mechanisms grows larger as the agents' (common) prior becomes more optimistic about the risky arm. We have the opposite result here. The set of implementable mechanisms is largest, albeit in a trivial fashion, for intermediate priors. For more extreme priors, this set is a singleton. This means that, but for intermediate priors, symmetric monotone mechanisms are always maximally transparent. These mechanisms are also always (weakly) less efficient than the perfect monitoring case, because they can never induce more experimentation, and sometimes strictly less.

Would private mechanisms help achieve outcomes that are more efficient than perfect monitoring equilibria?[76] We do not develop a formal proof, but it appears unlikely. The class of Markovian mechanisms allow the designer to conceal information from agents only when they are using the safe arm exclusively. Not disclosing information then means that these agents, conditional on no failure arriving, are more pessimistic about the risky arm than the designer is. Consequently, she is able to induce agents into using the safe arm for longer than they would have under perfect monitoring. Since the inefficiency in this model comes from under-experimentation relative to the cooperative benchmark, this suggests that private mechanisms would simply enable the designer to engineer even less efficient outcomes, not more.[77]

2.6 Discussion and Extensions

2.6.1 Non-Markovian Information Structures

Here we give a heuristic development of the claim in the introduction that allowing for punishments enables the designer to support a large set of equilibrium outcomes. Consider a discretisation of the basic game, as in Hörner et al. (2015), with news that

[76] Keller and Rady (2015) also consider asymmetric MPE, but they lead to ambiguous welfare comparisons with the symmetric MPE.

[77] This discussion assumes, as we have so far, that mechanisms are deterministic. Stochastic mechanisms in private signals (by theorem 2.7) may possess better efficiency properties. However, this introduces significant difficulty in the analysis, as outlined in Bonatti and Hörner (2017b).

is good but inconclusive. Let $\Delta > 0$ be the period length. Agents can adjust their actions, now restricted to $k_i \in \{0,1\}$, only at times $t = 0, \Delta, 2\Delta, \ldots$. Define $\delta = e^{-r\Delta}$.

Suppose one agent, whom we shall think of as a social planner, controlled all N bandits (along with observing their outcomes), but was constrained to choose the same action across all the arms.[78] Their Bellman equation, for the maximised discounted average payoff across all N arms, given a belief p, is

$$W\left(p\right) = \max\left\{s, \left(1-\delta\right)\lambda\left(p\right)h + \delta\mathbb{E}\left[W\left(p'\right)\right]\right\},$$

where p' is their posterior upon observing the outcome of N experiments. Let $\hat{p}$ be the cut-off belief that characterises the the optimal policy in the above optimisation problem, where the planner experiments at all beliefs above $\hat{p}$ and stops experimenting at any belief below $\hat{p}$.

Consider now the following mechanism. In each period, the designer first announces her belief p, and then recommends an action to each agent. If the designer's belief is above $\hat{p}$, then she recommends that all agents choose $k = 1$. Otherwise, she recommends $k = 0$. Deviators are punished with no information for the rest of the game.

It is clear that if all agents employ strategies that obey the designer's recommendations, their common value function is given by $W\left(p\right)$. Let us examine the decision problem of an agent, supposing that all other agents obey the designer's recommendations. If the agent solved the problem in autarky, then their optimal policy is characterised by a cut-off belief $\bar{p} > \hat{p}$, where they pull the risky arm whenever their belief is above $\bar{p}$ and pull the safe arm below $\bar{p}$. If the (public) belief p were greater than $\bar{p}$, it is easy to see that they will obey the designer's recommendation. However, if p is such that they would not be willing to experiment in autarky, then they follow the designer's recommendation if and only if

$$\left(1-\delta\right)\lambda\left(p\right)h + \delta\mathbb{E}\left[W\left(p'\right)\right] \geq s.$$

In particular, the agent's tradeoff is *exactly the same* as the tradeoff facing the designer: the designer recommends experimentation for those beliefs where the above inequality holds, and recommends the safe arm whenever it does not. Hence, the agent will obey the designer's recommendation. Moreover, the above argument easily extends to cut-offs that lie in $[\hat{p}, \bar{p}]$, by applying the above argument to the value function induced by the alternative choice of cut-off and making the observation that the induced value function is decreasing in the choice of cut-off on $[\hat{p}, \bar{p}]$.

[78] This constraint will bind, since the unconstrained optimal policy may involve (for some beliefs) using both risky and safe arms simultaneously across different bandits. However, the welfare loss due to the constraint disappears as $\Delta \to 0$. This is because Bellman equation for the unconstrained problem converges to the HJB equation in Keller and Rady (2010), where the optimal policy is extremal. See Hörner et al. (2015) for details.

2.6.2 Mixed Information Structures

We now allow the designer to use stochastic mechanisms.[79] We construct a simple stochastic mechanism where the designer randomises over a pair of symmetric monotone mechanisms. We consider only a setting with conclusive good news.

2.6.2.1 Random Cut-offs

Suppose that the designer randomises over a pair of symmetric monotone mechanisms. That is, at the start of the game, the designer randomly chooses between one of two cut-offs, $\tilde{q}_1$ and $\tilde{q}_2$ such that $p_N^* \leq \tilde{q}_1 < \tilde{q}_2 \leq p_1^*$. In particular, she chooses the cut-off $\tilde{q}_1$ with probability α, and the cut-off $\tilde{q}_2$ with probability $1-\alpha$.[80]

The HJB equation is the same as before, which means it has the solution (2.4.6). However, the mechanism above induces different boundary conditions. In particular, at $q = \tilde{q}_2$, we must have that

$$\bar{V}_i(p, \tilde{q}_2) = \left[\alpha + (1-\alpha)\frac{p-\tilde{q}_2}{1-\tilde{q}_2}\right]\underline{V}_i\left(\pi(p), \tilde{q}_2\right) + (1-\alpha)\left(\frac{1-p}{1-\tilde{q}_2}\right)s, \tag{2.6.1}$$

with

$$\pi(p) = \frac{\alpha p + (1-\alpha)\frac{p-\tilde{q}_2}{1-\tilde{q}_2}}{\alpha + (1-\alpha)\frac{p-\tilde{q}_2}{1-\tilde{q}_2}}.$$

To understand the expressions above, first notice that the denominator of $\pi(p)$ is the unconditional probability of the designer continuing to recommend experimentation at $q = \tilde{q}_2$, while the numerator is the joint probability that $\theta = 1$, and the designer continues to recommend experimentation. On the other hand, the right-hand side of (2.6.1) is just the expected continuation value of i in state $(p, \tilde{q}_2)$. With probability

$$\alpha + (1-\alpha)\frac{p-\tilde{q}_2}{1-\tilde{q}_2},$$

the designer continues to recommend experimentation, and i's continuation value is $\underline{V}_i\left(\pi(p), \tilde{q}_2\right)$, which is characterised below. With complementary probability, the designer recommends the safe arm, which gives a continuation value of s. At $q = \tilde{q}_1$,

$$\underline{V}_i(p, \tilde{q}_1) = \frac{1-p}{1-\tilde{q}_1}s + \frac{p-\tilde{q}_1}{1-\tilde{q}_1}\lambda_1 h.$$

This is exactly condition (2.4.6). Hence, for $q \in (\tilde{q}_1, \tilde{q}_2)$, i's value satisfies

$$\underline{V}_i(p, q) = p\lambda_1 h + (1-p)\frac{s - \tilde{q}_1\lambda_1 h}{1-\tilde{q}_1}\left[\frac{\Omega(q)}{\Omega(\tilde{q}_1)}\right]^{\frac{r}{N\lambda_1}},$$

[79] See footnote 39 on page 35 for the definition.
[80] We assume that p_0 is close enough to 1 so that the cut-off $\tilde{q}_2$ is implementable when $\alpha = 0$.

and the obedience constraint is the same as that shown in (2.4.9), *mutatis mutandis*. Note, however, that proposition 2.4.11 does not apply since the equality (2.4.10) no longer holds in this range.

These equations allow us to solve for the unknown function $\bar{F}$ that determines $\bar{V}_i$. In particular, we must have that

$$\bar{F}(x) = \left[\alpha\left(\frac{s-\tilde{q}_1\lambda_1 h}{1-\tilde{q}_1}\right)\left[\frac{\Omega(\tilde{q}_2)}{\Omega(\tilde{q}_1)}\right]^{\frac{r}{N\lambda_1}} + (1-\alpha)\left(\frac{s-\tilde{q}_2\lambda_1 h}{1-\tilde{q}_2}\right)\right]\left[x\Omega(\tilde{q}_2)\right]^{-\frac{r}{N\lambda_1}},$$

so that

$$\bar{V}_i(p,q) = p\lambda_1 h + (1-p)\,\bar{C}\left[\frac{\Omega(q)}{\Omega(\tilde{q}_2)}\right]^{\frac{r}{N\lambda_1}}.$$

Here,

$$\bar{C} = \alpha\left(\frac{s-\tilde{q}_1\lambda_1 h}{1-\tilde{q}_1}\right)\left[\frac{\Omega(\tilde{q}_2)}{\Omega(\tilde{q}_1)}\right]^{\frac{r}{N\lambda_1}} + (1-\alpha)\left(\frac{s-\tilde{q}_2\lambda_1 h}{1-\tilde{q}_2}\right).$$

Notice that when $\alpha = 0$, $\bar{V}_i$ simplifies to the value when there is a single cut-off given by $\tilde{q}_2$. It is straightforward to extend the arguments in section 2.4.4 to show that for fixed $\tilde{q}_1$, the pair of cut-offs $(\tilde{q}_1, \tilde{q}_2)$ is implementable for α close enough to 0. Conversely, for fixed α, the pair of cut-offs is implementable for $\tilde{q}_1$ close enough to $\tilde{q}_2$.

This extension demonstrates two points. First, the use of stochastic mechanisms allows the designer to partially reveal information to agents. To see this, notice that when $q = \tilde{q}_2$, and the designer continues to recommend experimentation, the agents infer that there is some probability that another agent has observed a success. This causes them to update their beliefs upward. However, because they know that with probability α, the designer would have continued to recommend experimentation in any history of the game, they are not certain that a success has arrived. Hence, the stochastic mechanism constructed above exhibits partial information disclosure. Second, stochastic mechanisms allow the designer to implement lower cut-offs than she otherwise would have been able to, albeit only probabilistically. In particular, note that even if the lowest implementable deterministic cut-off $\tilde{q}$ is greater than p_N^*, the designer can implement a more efficient outcome by randomising between $\tilde{q}$ and some other cut-off lower than $\tilde{q}$ in the fashion described above.

2.7 Conclusion

In this chapter, we analysed the problem of information disclosure in a setting of strategic experimentation. We focused on a particularly tractable class of mechanisms and studied their efficiency and transparency properties. It turned out that these properties depend heavily on the structure of the news arrival process.

This chapter still leaves many open questions. Are there appropriate restrictions on the class of information structures available to the designer that keep the problem tractable but non-trivial? What is the optimal information structure for the case when news is inconclusive, or when the state of the world changes over time? We hope to address these questions in future work.

2.8 Appendix A: Summary of Results from Keller et al. (2005); Keller and Rady (2010, 2015)

We summarise some useful results from the classic strategic experimentation game. $p^m = \frac{s}{\lambda_1 h}$ is the myopic cut-off belief, below which a myopic agent exclusively uses the safe arm, and above which exclusively uses the risky arm. Recall that $\Omega(p) = \frac{1-p}{p}$. The following is proposition 3.1 in Keller et al. (2005).

Proposition 2.2. *In the N-agent cooperative problem, there is a cut-off belief p_N^* given by*

$$p_N^* = \frac{rs}{(r + N\lambda_1)(\lambda_1 h - s) + rs}$$

such that below the cut-off it is optimal for all to play S exclusively and above it is optimal for all to play R exclusively. The value function V_N^* for the N-agent cooperative is given by

$$V_N^*(p) = p\lambda_1 h + (1-p)\frac{s - \lambda_1 p_N^* h}{1 - p_N^*}\left[\frac{\Omega(p)}{\Omega(p_N^*)}\right]^{\frac{r}{N\lambda_1}}$$

when $p > p_N^*$, and $V_N^*(p) = s$ otherwise.

Above the cut-off p_N^*, V_N^* satisfies the ordinary differential equation

$$N\lambda_1 p(1-p)u'(p) + (r + N\lambda_1 p)u(p) = (r + N\lambda_1)p\lambda_1 h,$$

which has the general solution

$$V_N(p) = p\lambda_1 h + C(1-p)\left[\Omega(p)\right]^{\frac{r}{N\lambda_1}}. \tag{2.8.1}$$

We also state their proposition 5.1.

Proposition 2.3. *The N-player experimentation game has a unique symmetric equilibrium in Markovian strategies with the common posterior belief as the state variable. In this equilibrium, the safe arm is used exclusively at beliefs below the single-player cut-off p_1^*; the risky arm is used exclusively at all beliefs above a cut-off $p_N^\dagger > p_1^*$ solving*

$$(N-1)\left(\frac{1}{\Omega(p^m)} - \frac{1}{\Omega\left(p_N^\dagger\right)}\right) = \left(1 + \frac{r}{\lambda_1}\right)\left[\frac{1}{1-p_N^\dagger} - \frac{1}{1-p_1^*} - \frac{1}{\Omega(p_1^*)}\ln\left(\frac{\Omega(p_1^*)}{\Omega\left(p_N^\dagger\right)}\right)\right];$$

and a positive fraction of the resource is allocated to each arm at beliefs strictly between p_1^* *and* $p_N^\dagger$*. The fraction of the resource that each player allocates to the risky arm at such a belief is*

$$k_N^\dagger(p) = \frac{1}{N-1}\frac{W^\dagger(p) - s}{s - p\lambda_1 h}$$

with

$$W^\dagger(p) = s + \frac{rs}{\lambda_1}\left[\Omega(p_1^*)\left(1 - \frac{1-p}{1-p_1^*}\right) - (1-p)\ln\left(\frac{\Omega(p_1^*)}{\Omega(p)}\right)\right],$$

which is each player's value function on $\left[p_1^*, p_N^\dagger\right]$ *and satisfies* $W^\dagger(p_1^*) = s$*,* $\left(W^\dagger\right)'(p_1^*) = 0$*. Below* p_1^* *the value function equals* s*, and above* $p_N^\dagger$ *it is given by* $V_N(p)$ *from equation (2.8.1) with* $V_N\left(p_N^\dagger\right) = W^\dagger\left(p_N^\dagger\right)$*.*

On the interval $\left[p_1^*, p_N^\dagger\right]$, $W^\dagger(p)$ solves the ODE

$$\lambda_1 p(1-p)u'(p) + \lambda_1 p u(p) = (r+\lambda_1)p\lambda_1 h - rs. \tag{2.8.2}$$

2.9 Appendix B: Omitted Proofs and Results

2.9.1 Derivation of Hamilton-Jacobi-Bellman Equations

We derive the HJB equation (2.4.5) for agent i when their belief does not coincide with the designer's. By the dynamic programming principle, agent i's value function satisfies

$$u(p,q) = \max_{k\in[0,1]}\left\{r\left[(1-k)s + kp\lambda_1 h\right] + e^{-r\mathrm{d}t}\mathbb{E}\left[u(p+\mathrm{d}p, q+\mathrm{d}q)\,|p,q,k\right]\right\}.$$

With probability $kp\lambda_1\mathrm{d}t$, a success arrives and the value jumps to $u(1,1) = \lambda_1 h$. With probability $1 - kp\lambda_1\mathrm{d}t$, no success arrives, and the value function changes to $u(p,q) + \partial_p u(p,q)\,\mathrm{d}p + \partial_q u(p,q)\,\mathrm{d}q$. The latter can, using (2.4.1) and (2.4.2), be written as

$$u(p,q) - k\lambda_1 p(1-p)\partial_p u(p,q)\,\mathrm{d}t - \lambda_1 q(1-q)\left[k + \sum_{j\neq i}\rho_j(q)\right]\partial_q u(p,q)\,\mathrm{d}t.$$

Replacing $e^{-r\mathrm{d}t}$ with $1 - r\mathrm{d}t$ and plugging in the above expectation (ignore terms of $O\left(\mathrm{d}t^2\right)$) gives equation (2.4.5).

2.9.2 Deriving Equation (2.4.10)

Integrating (2.4.1) along the conjectured equilibrium, and using the initial condition $q_0 = p_0$, yields

$$q_t = \frac{p_0 e^{-N\lambda_1 t}}{p_0 e^{-N\lambda_1 t} + 1 - p_0}.$$

Notice now that

$$\frac{\Omega\left(q_t\right)}{\Omega\left(p_0\right)} = e^{N\lambda_1 t} = \left[\frac{\Omega\left(p_t\right)}{\Omega\left(p_0\right)}\right]^N ,$$

where the last equality can be found by integrating (2.4.2) in the same way. Note that the chain of equalities above can also be used to derive (2.4.7) directly from first principles.

2.9.3 Proof of Proposition 2.1

We first establish the comparative statics on p. $MB\left(p;\tilde{q}\right)$, given in (2.4.11), can be written as

$$p\lambda_1 h + C\left(1-p\right)\Omega\left(p\right)^{\frac{r}{\lambda_1}} , \tag{2.9.1}$$

where C is a positive constant. Differentiating (2.9.1), we have

$$\lambda_1 h - C\left[1+\frac{r}{\lambda_1 p}\right]\Omega\left(p\right)^{\frac{r}{\lambda_1}} .$$

This expression goes to $-\infty$ as $p\downarrow 0$, and is positive when $p=1$. The second derivative of (2.9.1) is

$$C\left[\frac{r}{\lambda_1 p\left(1-p\right)}\right]\left[\frac{\lambda_1+r}{\lambda_1 p}\right]\left[\Omega\left(p\right)\right]^{\frac{r}{\lambda_1}} > 0,$$

establishing the strict convexity of $MB\left(p;\tilde{q}\right)$ in p.

To see that C is positive, notice that the sign of C depends on $W^\dagger\left(\tilde{q}\right)-\tilde{q}\lambda_1 h$. If $\tilde{q}<p_1^*$, then, from proposition 2.3, $W^\dagger\left(\tilde{q}\right)=s$. Proposition 2.2 then guarantees that $W^\dagger\left(\tilde{q}\right)-\tilde{q}\lambda_1 h = s-\tilde{q}\lambda_1 h > 0$. If $\tilde{q}\in\left[p_1^*,p_N^\dagger\right]$, then proposition 2.3 gives us that $W^\dagger\left(\tilde{q}\right) = s+\left(N-1\right)\left(s-\tilde{q}\lambda_1 h\right)k^\dagger\left(\tilde{q}\right)$, so that $W^\dagger\left(\tilde{q}\right)-\tilde{q}\lambda_1 h = \left[\left(N-1\right)k^\dagger\left(\tilde{q}\right)+1\right]\left(s-\tilde{q}\lambda_1 h\right)$. $k^\dagger$ is bounded below by 0, and $s-\tilde{q}\lambda_1 h>0$ also by proposition 2.3. Finally, if $\tilde{q}>p_N^\dagger$, then

$$W^\dagger\left(\tilde{q}\right)-\tilde{q}\lambda_1 h = \left(1-\tilde{q}\right)\frac{W^\dagger\left(p_N^\dagger\right)-p_N^\dagger\lambda_1 h}{1-p_N^\dagger}\left[\frac{\Omega\left(\tilde{q}\right)}{\Omega\left(p_N^\dagger\right)}\right]^{\frac{r}{N\lambda_1}} > 0,$$

where the inequality follows in a similar fashion.

We now establish the second part of the proposition. First note that $MB\left(p;\tilde{q}\right)$ can, ignoring terms that do not depend on $\tilde{q}$, be written as

$$\hat{C}\frac{W^\dagger\left(\tilde{q}\right)-\tilde{q}\lambda_1 h}{1-\tilde{q}}\left[\Omega\left(\tilde{q}\right)\right]^{-\frac{r}{N\lambda_1}} ,$$

where again $\hat{C}$ is a positive constant. The derivative of the above expression with respect to $\tilde{q}$ thus has the same sign as

$$\frac{1}{N\lambda_1\tilde{q}\left(1-\tilde{q}\right)^2\left[\Omega\left(\tilde{q}\right)\right]^{\frac{r}{N\lambda_1}}}\left[N\lambda_1\tilde{q}\left(1-\tilde{q}\right)\left(\left(W^{\dagger}\right)'\left(\tilde{q}\right)-\lambda_1 h\right)+\left(r+N\lambda_1\tilde{q}\right)\left(W^{\dagger}\left(\tilde{q}\right)-\tilde{q}\lambda_1 h\right)\right], \tag{2.9.2}$$

whose sign clearly depends only on the expression in brackets. Suppose $\tilde{q} \in \left[p_1^*, p_N^{\dagger}\right)$. From (2.8.2) and proposition 2.3, we can write the expression in brackets as

$$r\left(N-1\right)\left(s-\tilde{q}\lambda_1 h\right)\left(k^{\dagger}\left(\tilde{q}\right)-1\right)<0,$$

where the inequality is deduced from the fact that, on $\left[p_1^*, p_N^{\dagger}\right)$, $k^{\dagger}\left(p\right)<1$. Suppose instead that $\tilde{q}<p_1^*$. Proposition 2.3 allows us to write the expression in brackets in (2.9.2) as

$$rs-\tilde{q}\left[N\lambda_1\left(\lambda_1 h-s\right)+r\lambda_1 h\right] \gtreqless 0 \iff \tilde{q} \lesseqgtr p_N^*.$$

The equivalence displayed above comes from proposition 2.2. This establishes the result.

2.9.4 Symmetric Monotone Mechanisms: Necessary and Sufficient Conditions

In this section we give an implicit but complete characterisation of the set of implementable cut-offs $\tilde{q}$. We first state an obvious sufficient condition for the implementability of $\tilde{q}$. It follows immediately from proposition 2.1, the envelope theorem and continuity.

Proposition 2.4. *Let $\underline{p}\left(\tilde{q}\right)$ attain the minimum (over p) of $MB\left(p;\tilde{q}\right)$. This minimum must be in the interior of the unit interval. If $MB\left(\underline{p}\left(\tilde{q}\right);\tilde{q}\right)\geq s$, then $\tilde{q}$ is implementable. Suppose $MB\left(\underline{p}\left(\tilde{q}\right);\tilde{q}\right)\geq s$. If $\underline{p}<p_N^*$, then every $q'\in\left[\tilde{q},p_N^*+\varepsilon\right)$, for some $\varepsilon>0$, is implementable as well. If $\underline{p}>p_N^*$, then the same holds for $q'\in\left(p_N^*-\varepsilon,\tilde{q}\right]$.*

Unfortunately, there is no simple way to guarantee that $MB\left(\underline{p};\tilde{q}\right)\geq s$. We thus state an alternative sufficient condition, which follows from the strict convexity and continuous differentiability of $MB\left(p;\tilde{q}\right)$. First, define the function $\hat{p}\left(q\right)$ by

$$\left[\frac{\Omega\left(\hat{p}\left(q\right)\right)}{\Omega\left(p_0\right)}\right]^N=\frac{\Omega\left(q\right)}{\Omega\left(p_0\right)}. \tag{2.9.3}$$

Proposition 2.5. *Let $\tilde{p}$ be the agents' common private belief under obedience, before the designer's true belief is revealed, when $q=\tilde{q}$. That is, $\tilde{p}=\hat{p}\left(\tilde{q}\right)$. If $MB\left(\tilde{p};\tilde{q}\right)\geq s$, and*

$$\frac{\partial MB}{\partial p}\left(\tilde{p};\tilde{q}\right)\geq 0,$$

then $\tilde{q}$ is implementable. Moreover, if both of the above inequalities are strict, then there exists an $\varepsilon>0$ such that if $q'\in\left(\tilde{q}-\varepsilon,\tilde{q}+\varepsilon\right)\cap\left[0,p_N^{\dagger}\right]$, then q' is implementable as well.

This proposition is useful because factors involving exponents disappear when $q = \tilde{q}$, which makes the conditions easier to verify. The next result is easy to see from the two propositions, and the strict convexity of $MB\left(p;\tilde{q}\right)$ in p.

Theorem 2.8. *If $\tilde{q}$ is implementable, then it must satisfy at least one of the sufficient conditions stated in propositions 2.4 and 2.5. In other words, propositions 2.4 and 2.5 collectively give necessary and sufficient conditions for the implementability of any cut-off $\tilde{q}$.*

We state a limited monotonicity result.

Proposition 2.6. *If $\tilde{q}$ satisfies the sufficient conditions in proposition 2.5, then every $q' \in \left[\tilde{q}, p_N^\dagger\right]$ that satisfies $MB\left(\hat{p}\left(q'\right);q'\right) \geq s$ is implementable.*

$MB\left(\hat{p}\left(q\right);q\right)$ is in general, as one would expect given proposition 2.1, not monotonic, which explains the final condition above.[81] To prove the result, we use a lemma, from which, when combined with proposition 2.5, the rest is immediate. Remember that $\tilde{p}$ is defined implicitly in terms of $\tilde{q}$.

Lemma 2.6.

$$\frac{\partial MB}{\partial p}\left(\hat{p}\left(\tilde{q}\right);\tilde{q}\right) \tag{2.9.4}$$

is increasing in $\tilde{q}$.

Proof. The derivative of (2.9.4) with respect to $\tilde{q}$ is

$$-\frac{1}{N}\left[1+\frac{r}{\lambda_1\hat{p}}\right]\underbrace{\left[\frac{(1-\tilde{q})\left[\left(W^\dagger\right)'(\tilde{q})-\lambda_1 h\right]+W^\dagger(\tilde{q})-\tilde{q}\lambda_1 h}{(1-\tilde{q})^2}\right]}_{(-)}+\underbrace{\frac{1}{N}\left[\frac{W^\dagger(\tilde{q})-\tilde{q}\lambda_1 h}{1-\tilde{q}}\right]\frac{r}{\lambda_1\hat{p}^2}\frac{\mathrm{d}\hat{p}}{\mathrm{d}\tilde{q}}}_{(+)}.$$

The fact that the second term is positive follows from the proof of proposition 2.1, which guarantees that the factor in brackets is positive, and from the implicit function theorem:

$$\frac{\mathrm{d}\hat{p}}{\mathrm{d}\tilde{q}}=\frac{\hat{p}\left(1-\hat{p}\right)}{N\tilde{q}\left(1-\tilde{q}\right)}>0.$$

As a consequence, it suffices to sign the first term, which we will show is also positive. This is an immediate consequence of the benefit to experimentation always being non-negative, as established in Keller et al. (2005). A direct proof, which we sketch, uses the same arguments as in the proof of proposition 2.1.

Suppose $\tilde{q} \leq p_1^*$. Then $(1-\tilde{q})\left[\left(W^\dagger\right)'(\tilde{q})-\lambda_1 h\right]+W^\dagger(\tilde{q})-\tilde{q}\lambda_1 h = s-\lambda_1 h < 0$. Suppose instead that $\tilde{q} > p_1^*$.[82] Then, by (2.8.2), $(1-\tilde{q})\left[\left(W^\dagger\right)'(\tilde{q})-\lambda_1 h\right]+W^\dagger(\tilde{q})-$

[81] See, however, the result after the lemma.
[82] Recall that corollary 2.3 requires that $\tilde{q} \leq p_N^\dagger$.

$\tilde{q}\lambda_1 h$ becomes

$$\frac{(r+\lambda_1)\,\tilde{q}\lambda_1 h - rs}{\lambda_1\tilde{q}} - \lambda_1 h = r\left(h - \frac{s}{\tilde{q}\lambda_1}\right) < 0.$$

□

We now state a simple sufficient condition that guarantees that $MB\left(\hat{p}(q);q\right) \geq s$, given that $\tilde{q} \leq q$ satisfies $MB\left(\hat{p}(\tilde{q});\tilde{q}\right) \geq s$.

Proposition 2.7. *Suppose $MB\left(\hat{p}(\tilde{q});\tilde{q}\right) \geq s$. If $\tilde{q} < p_1^*$, then $MB\left(\hat{p}(q);q\right) \geq s$ for all $q \in \left[\tilde{q}, p_1^*\right]$. If $\tilde{q} \geq p_1^*$, then there exists r_0 such that for all $r \in (0, r_0)$, $MB\left(\hat{p}(q);q\right) \geq s$ for all $q \in \left[\tilde{q}, p_N^{\dagger}\right]$.*

Proof. Suppose $MB\left(\hat{p}(\tilde{q});\tilde{q}\right) \geq s$. Then, we must have

$$\frac{1}{N}\left(\frac{W^{\dagger}(\tilde{q}) - \tilde{q}\lambda_1 h}{1-\tilde{q}}\right) \geq \frac{s - \hat{p}(\tilde{q})\,\lambda_1 h}{1-\hat{p}(\tilde{q})}.$$

The derivative of the left-hand expression with respect to $\tilde{q}$ is

$$\frac{1}{N}\left[\frac{(1-\tilde{q})\left[\left(W^{\dagger}\right)'(\tilde{q}) - \lambda_1 h\right] + W^{\dagger}(\tilde{q}) - \tilde{q}\lambda_1 h}{(1-\tilde{q})^2}\right],$$

while the derivative of the right-hand expression is

$$\frac{\hat{p}(\tilde{q})\,(s-\lambda_1 h)}{N\tilde{q}\,(1-\tilde{q})\left[1-\hat{p}(\tilde{q})\right]}.$$

Thus, our inequality remains to be true after an increase in $\tilde{q}$ if and only if

$$\left[\frac{(1-\tilde{q})\left[\left(W^{\dagger}\right)'(\tilde{q}) - \lambda_1 h\right] + W^{\dagger}(\tilde{q}) - \tilde{q}\lambda_1 h}{1-\tilde{q}}\right] \geq \frac{\hat{p}(\tilde{q})\,(s-\lambda_1 h)}{\tilde{q}\left[1-\hat{p}(\tilde{q})\right]}.$$

For $\tilde{q} \leq p_1^*$, this inequality becomes

$$\frac{s-\lambda_1 h}{1-\tilde{q}} \geq \frac{\hat{p}(\tilde{q})\,(s-\lambda_1 h)}{\tilde{q}\left[1-\hat{p}(\tilde{q})\right]},$$

which is equivalent to

$$\Omega\left(\hat{p}(\tilde{q})\right) \leq \Omega(\tilde{q}).$$

This is readily verified to be true. If $\tilde{q} > p_1^*$, then we have

$$r\left[\frac{\tilde{q}\lambda_1 h - s}{\lambda_1\tilde{q}\,(1-\tilde{q})}\right] \geq \frac{\hat{p}(\tilde{q})\,(s-\lambda_1 h)}{\tilde{q}\left[1-\hat{p}(\tilde{q})\right]},$$

which can be written as

$$\frac{r}{\lambda_1}\left(\tilde{q}\lambda_1 h - s\right)\Omega\left(\hat{p}\left(\tilde{q}\right)\right) \geq \left(1-\tilde{q}\right)\left(s-\lambda_1 h\right).$$

We wish for this inequality to be true for all small enough r uniformly over $\tilde{q} \in \left(p_1^*, p_N^\dagger\right)$. In view of the fact that the left-hand side is decreasing while the right-hand side is increasing, we must establish that

$$\frac{r}{\lambda_1}\left(p_1^*\lambda_1 h - s\right)\Omega\left(\hat{p}\left(p_1^*\right)\right) \geq \left(1-p_N^\dagger\right)\left(s-\lambda_1 h\right),$$

so that our choice of $\tilde{q}$ is irrelevant. As $r \downarrow 0$, the right-hand side converges to a finite negative number.[83] However, the left-hand side, after an application of L'Hôpital's rule, converges to 0.[84] □

The above proposition does not yet allow us to conclude that the set of implementable cut-offs has an interval structure, since proposition 2.6 only applies when the conditions given in proposition 2.5 hold. We therefore need to establish when this is the case, which is done below.

Proposition 2.8. *If $\tilde{q} \in \left[p_N^*, p_N^\dagger\right]$ is implementable, then, for all r close enough to zero,*

$$\frac{\partial MB}{\partial p}\left(\hat{p}\left(\tilde{q}\right);\tilde{q}\right) \geq 0.$$

This condition is equivalent to $\underline{p}\left(\tilde{q}\right) \leq \hat{p}\left(\tilde{q}\right)$.

Proof. We will prove the latter statement, whose equivalence is due to the strict convexity of $MB\left(p;\tilde{q}\right)$. That is, we wish to prove that for $\tilde{q} \in \left[p_N^*, p_N^\dagger\right]$ and r close enough to zero, $\underline{p}\left(\tilde{q}\right) \leq \hat{p}\left(\tilde{q}\right)$. Since $\tilde{q} \leq \hat{p}\left(\tilde{q}\right)$, it will be sufficient to show that $\underline{p}\left(\tilde{q}\right) \leq \tilde{q}$. Note that $\underline{p}\left(\tilde{q}\right)$ solves the first order condition

$$\lambda_1 h = \frac{1}{N}\left[1+\frac{r}{\lambda_1 \underline{p}}\right]\left[\frac{\Omega\left(\underline{p}\right)}{\Omega\left(p_0\right)}\right]^{\frac{r}{\lambda_1}}\left[\frac{W^\dagger\left(\tilde{q}\right)-\tilde{q}\lambda_1 h}{1-\tilde{q}}\right]\left[\frac{\Omega\left(p_0\right)}{\Omega\left(\tilde{q}\right)}\right]^{\frac{r}{N\lambda_1}}.$$

This solution must be unique by strict convexity of $MB\left(p;\tilde{q}\right)$. First observe that the right-hand side above is decreasing in $\underline{p}$. As a consequence, we will demonstrate that

$$\lambda_1 h \geq \frac{1}{N}\left[1+\frac{r}{\lambda_1 p}\right]\left[\frac{\Omega\left(p\right)}{\Omega\left(p_0\right)}\right]^{\frac{r}{\lambda_1}}\left[\frac{W^\dagger\left(\tilde{q}\right)-\tilde{q}\lambda_1 h}{1-\tilde{q}}\right]\left[\frac{\Omega\left(p_0\right)}{\Omega\left(\tilde{q}\right)}\right]^{\frac{r}{N\lambda_1}}$$

when p is evaluated at $\tilde{q}$. Recall that

$$\left[\frac{W^\dagger\left(\tilde{q}\right)-\tilde{q}\lambda_1 h}{1-\tilde{q}}\right]\left[\frac{\Omega\left(p_0\right)}{\Omega\left(\tilde{q}\right)}\right]^{\frac{r}{N\lambda_1}}$$

[83] In particular, $p_N^\dagger \to \Omega^{-1}\left(\frac{N}{N-1}\Omega\left(p^m\right)\right)$.

[84] Recall that $p_1^* \to 0$ as $r \to 0$. We then deduce that $\frac{r}{\hat{p}\left(p_1^*\right)} \to 0$.

is decreasing in $\tilde{q}$ on $\left(p_N^*, p_N^\dagger\right)$. Hence, it will be enough to prove that

$$\begin{aligned}\lambda_1 h &\geq \frac{1}{N}\left[1+\frac{r}{\lambda_1 p}\right]\left[\frac{\Omega(p)}{\Omega(p_0)}\right]^{\frac{r}{\lambda_1}}\left[\frac{W^\dagger\left(p_N^*\right)-p_N^*\lambda_1 h}{1-p_N^*}\right]\left[\frac{\Omega(p_0)}{\Omega\left(p_N^*\right)}\right]^{\frac{r}{N\lambda_1}} \\ &= \frac{1}{N}\left[1+\frac{r}{\lambda_1 p}\right]\left[\frac{\Omega(p)}{\Omega(p_0)}\right]^{\frac{r}{\lambda_1}}\left[\frac{s-p_N^*\lambda_1 h}{1-p_N^*}\right]\left[\frac{\Omega(p_0)}{\Omega\left(p_N^*\right)}\right]^{\frac{r}{N\lambda_1}}.\end{aligned}$$

Letting $r \downarrow 0$, so that $p_N^* \to 0$, and applying L'Hôpital's rule, recovers the strict inequality $\lambda_1 h > 0$, proving the result. $\square$

We thus have enough to prove theorem 2.3. This is done in appendix 2.9.6.

2.9.5 Proof of Theorem 2.2

We first prove that the cut-off $\tilde{q} = p_N^\dagger$ is implementable. If $p_0 < p_N^\dagger$, then the designer reveals all information immediately and recommends equilibrium actions by construction. Suppose instead that $p_0 \geq p_N^\dagger$. We shall make use of proposition 2.5. A rearrangement of the inequality $MB(p;\tilde{q}) \geq s$ yields

$$\frac{1}{N}\frac{W^\dagger(\tilde{q})-\tilde{q}\lambda_1 h}{1-\tilde{q}} \geq \frac{s-p\lambda_1 h}{1-p}.$$

Let $\tilde{q} = p_N^\dagger$ and $p = \hat{p}\left(p_N^\dagger\right)$.[85] We know from proposition 2.3 that $W^\dagger\left(p_N^\dagger\right) - p_N^\dagger\lambda_1 h = N\left(s - p_N^\dagger\lambda_1 h\right)$. It is also easy to see that $\hat{p}(q) \geq q$, with strict inequality when $q < p_0$. As a consequence, it suffices to show that

$$\frac{s-p\lambda_1 h}{1-p}$$

is strictly decreasing. Inspecting the derivative of the above expression shows this to be the case. We now have that $MB\left(\hat{p}\left(p_N^\dagger\right); p_N^\dagger\right) \geq s$, with strict inequality when $p_N^\dagger < p_0$. It therefore remains to be established that

$$\lambda_1 h - \frac{1}{N}\left[1+\frac{r}{\hat{p}\lambda_1}\right]\left[\frac{W^\dagger\left(p_N^\dagger\right)-p_N^\dagger\lambda_1 h}{1-p_N^\dagger}\right] \geq 0.$$

Using the same arguments as above, the inequality can be written as

$$\lambda_1 h\left[\frac{\hat{p}\lambda_1}{\hat{p}\lambda_1 + r}\right] \geq \frac{s-p_N^\dagger\lambda_1 h}{1-p_N^\dagger}.$$

[85] See (2.9.3) for the definition of $\hat{p}$.

Notice that the left hand side can be bounded below as shown:

$$\lambda_1 h \left[\frac{\hat{p}\lambda_1}{\hat{p}\lambda_1 + r}\right] \geq \lambda_1 h \left[\frac{p_N^{\dagger}\lambda_1}{p_N^{\dagger}\lambda_1 + r}\right],$$

where the inequality is strict whenever $p_N^{\dagger} < p_0$. Hence, it is enough to prove that

$$\lambda_1 h \left[\frac{p_N^{\dagger}\lambda_1}{p_N^{\dagger}\lambda_1 + r}\right] \geq \frac{s - p_N^{\dagger}\lambda_1 h}{1 - p_N^{\dagger}}.$$

Simplifying the above expression, and replacing $p_N^{\dagger}$ with p, we have

$$p \geq \frac{rs}{\lambda_1 (\lambda_1 h - s) + r\lambda_1 h},$$

which is strictly true for $p = p_N^{\dagger}$ via propositions 2.2 and 2.3.

2.9.6 Proof of Theorem 2.3

Suppose that $\tilde{q} \geq p_N^*$, and fix some $q' > \tilde{q}$. Since $\tilde{q}$ is implementable, we must have that $MB\left(\hat{p}\left(\tilde{q}\right);\tilde{q}\right) \geq s$. We therefore have, from proposition 2.7, that, for an appropriate choice of r, $MB\left(\hat{p}\left(q'\right);q'\right) \geq s$. Proposition 2.8 also guarantees that

$$\frac{\partial MB}{\partial p}\left(\hat{p}\left(\tilde{q}\right);\tilde{q}\right) \geq 0,$$

which is increasing in $\tilde{q}$, by 2.6. The cut-off q' therefore satisfies the sufficient conditions in proposition 2.5, which guarantees implementability.

Suppose instead that $\tilde{q} < p_N^*$. By corollary 2.4, every $q' \in \left[\tilde{q}, p_N^*\right]$ is also implementable. Applying the previous argument for $q' = p_N^*$ gives the result.

3 Social Learning and the Persistent Effects of Institutions

3.1 Introduction

There is a large and growing literature demonstrating that historical institutions have had significant impacts on contemporary economic outcomes.[1] Banerjee and Iyer (2005) showed that differences in historical property rights institutions in India lead to sustained differences in levels of agricultural investment and productivity, and investment in health and education. Nunn (2008) found a robust negative relationship between the number of slaves exported from African countries and the current economic performance of the countries they were exported from. The poorest countries in Africa today are those from which the largest numbers of slaves were exported. Dell (2010) examined the long-run impacts of the *mita*, an extensive forced mining labour system imposed in Peru and Bolivia between 1573 and 1812. She found that previous exposure to the *mita* lowered present household consumption by 25% and increased the current prevalence of stunted growth in children by about 6%. The *mita* also lowered education and agricultural market participation. The common thread throughout this body of research is that bad, extractive institutions lead to persistently adverse economic outcomes.

These results highlight just how important these historical institutions are, and the importance of institutions drives research in this area. However, the growth of empirical work has far outpaced the growth of theoretical work. This paper seeks to fill that gap.

We do not have any well-known formal theories to explain why institutions in general might have persistent effects.[2] One reason why this is the case might be that there just is no broad theory to explain this phenomenon; there are only explanations that are specific to each instance and vary across different kinds of institutions. Another reason is that institutions are difficult to define[3], which makes it challenging to formalise any

[1] See, for example, Acemoglu et al. (2001, 2002); Glaeser and Shleifer (2002); Acemoglu and Johnson (2005); Banerjee and Iyer (2005); Field (2007); Nunn (2008); Galor et al. (2009); Nunn and Wantchekon (2011); Dell (2010); Greif and Tabellini (2010); Voigtländer and Voth (2012); Jha (2013); Michalopoulos and Papaioannou (2013); Acemoglu and Robinson (2012). Nunn (2009) has a good survey of the literature.

[2] North's framework (2006) of economic agents with 'mental models' is probably closest to the model we develop. However, he does not formalise this idea.

[3] Greif and Kingston (2011) consider various conceptions of institutions and their theoretical implications. Greif and Laitin (2004); Roland (2004); Kingston and Caballero (2009) each cover various

general theories about their effects. We hope to overcome these issues by borrowing insights from social learning theory.[4] Roughly speaking, we posit that institutions influence behaviour in a way that makes that behaviour persistent. Hence, even when institutions are reformed, the effects associated with those institutions persist, because the behaviour associated with the institutions still persists. The linchpin of this argument is then the mechanism that makes behaviour persistent. Social learning provides this mechanism.

We adapt the canonical social learning model by Bikhchandani et al. (1992), hereafter referred to as BHW.[5]The key insight in BHW is that when you observe a long enough history of actions that precede you, at a certain point, it becomes optimal to ignore any private information you have and instead rely solely on the information contained in that history. This phenomenon is called an *informational cascade*. In a sense, a long enough history outweighs the information any one individual might have. When this happens, individuals start to herd on the action that history suggests to be optimal. Moreover, since each individual is ignoring their own private information and simply following the herd, there is no new private information that is being made publicly available. This is because actions in a herd have no informational content. Hence the publicly known information never changes, so the action it prescribes does not change as well. As a consequence, once herds start, they never end. What makes the BHW result truly remarkable, though, is that there is a positive probability of herding on a wrong outcome.

In BHW, payoffs are determined by some unknown and unchanging state of the world, and the optimal action depends on which state one is in. Here, as we develop later, payoffs are determined by the interaction of an unknown state of the world and the observable quality of institutions that one is exposed to. By abstracting from different kinds of institutions and instead focusing on whether their effects are (relatively) good or (relatively) bad, we sidestep the difficulty in defining institutions. Hence, when we say "institutions", we actually mean "effects of institutions" or "quality of institutions", but simply use the former to avoid being overly wordy. It is this interaction that creates herd behaviour in our setup. Certain actions can be explained by either extractive institutions, negative information about the state of the world, or both. When extractive institutions persist for a long enough time, they eventually induce pessimistic beliefs about the state of the world, and this creates herding on low-payoff outcomes.

We propose that the contemporary underdevelopment causally linked with extractive

theories of how institutions, broadly conceived, change. The conceptual issues they face in doing so appear to stem from difficulties in pinning down exactly what institutions are. Acemoglu and Robinson (2012) also invoke the idea of institutions in a very informal sense.

[4]Social learning has been used to explain a wide variety of phenomena such as technology adoption. See, for instance, Bandiera and Rasul (2006); Conley and Udry (2010). For a survey of theoretical results in this field, see Smith and Sørensen (2011); Moscarini and Smith (1997). For a textbook-length treatment, see Chamley (2004).

[5]Banerjee (1992) develops a similar theory of herd behaviour. The approach in BHW better suits our purposes.

institutions in the literature can be thought of as individuals herding on actions that lead to poor outcomes. Herding on a poor outcome may have been optimal under extractive institutions, but is no longer the correct course of action after institutions have been reformed. For example, it is probably optimal for someone to exert low effort when the state extracts most, if not all, of the fruits of that person's labour. Under some conditions, these herds will continue to persist despite reforms that make institutions less extractive. As a consequence, to continue the previous example, individuals will continue to exert low effort even if it is now more rewarding to exert high effort instead, and it is the continued herding on the low-effort action that drives underdevelopment.

In the model developed later, persistent herding on bad outcomes is possible only under two conditions. One is when institutions are degenerate, which obtains when institutions do not affect individual choices whenever there is no uncertainty about the state of the world. Hence, when institutions are degenerate, they only have impacts on decisions in so far as one is uncertain about the state of the world. The other condition is when institutions and states of the world are complements: one needs both institutions and states to be favourable for high effort to pay off. In both cases, the prior belief on the state of the world must not be too optimistic, although less so under complementarity than degeneracy. On the other hand, when good institutions are substitutes for a favourable state of the world, herding on bad outcomes is impossible. To be sure, there is a large class of models that can form the basis of a theory of short-lived institutions with long-lived effects. Contained in this class are models exhibiting path-dependence, such as those in Fernandez and Rodrik (1991); Krusell and Rios-Rull (1996); Coate and Morris (1999); Acemoglu and Jackson (2015), whose ideas can be used to develop competing theories of persistence. There are a few reasons why we focus on the subclass of models that involve social learning.

First, social learning presents a persuasive account of how people actually behave. The literature in psychology documents a phenomenon called *informational social influence* (Cialdini and Goldstein, 2004; Deutsch and Gerard, 1955), wherein people tend to look to others for information on how to act in situations of uncertainty. Economic social learning is just the formal counterpart of this and thusly is, at least partly, supported by the relevant literature in psychology in a way that competing theories will not be.

Second, as we will see later, social learning provides us with a general theory that is applicable to a wide array of institutions, because it allows us to abstract from the specific forms that institutions take and forces us to focus on their effects instead. Other models about the persistence of institutions, such as in Piketty (1995); Acemoglu and Robinson (2001); Benabou and Ok (2001); Acemoglu (2003); Bénabou and Tirole (2006); Padró i Miquel (2007); Nunn (2007); Acemoglu and Robinson (2008), tend to be only about specific institutions and hence have explanatory power limited only to the specific types of institutions they consider. For instance, Piketty (1995) examines redistributive institutions and shows that persistent differences in redistribution regimes can be traced to individual mobility experience, which is itself affected by beliefs about processes of mobility in a self-fulfilling way. Nunn (2007) builds a model with production externalities, which results in an economy trapped in a low-production equilibrium when exposed to extractive colonisers. Acemoglu and Robinson (2008) considers how elites trade off between formal political and informal economic power when they influence transitions between democracy and non-democracy. These models are only readily applicable to the institutions that they specify.

Third, generality notwithstanding, social learning gives us a model that is both consistent with the empirical findings about institutions, but also yields some theoretical predictions, as described above, that make this theory, or, at least, a richer version of it, testable.[6]

Finally, it also gives us a model that is simple and tractable, as evidenced by the relative ease with which many of the early results are obtained. The model by BHW has many advantages, at least some of which have just been outlined. Predictably, it is not without its weaknesses. In particular, the herding-on-a-wrong-outcome result is not robust to changes in its assumptions about the discreteness of the action space and signal space (Lee, 1993; Smith and Sørensen, 2000, 2011; Smith et al., 2017). We attempt to address this criticism in section 3.3.

This chapter proceeds as follows. In the next section, we present the basic model and some results. In section 3.3, we develop two extensions of the basic model. Section 3.4 addresses the criticism on the robustness of wrong herds and considers the applicability of the theory to empirical findings. Section 3.5 concludes.

3.2 Basic Model

3.2.1 Assumptions

Suppose there is an infinite sequence of agents ordered exogenously and indexed by $t \in \mathbb{N}$. t is also the time period that each agent acts. Each agent acts only once.[7]

Preferences are defined over a random outcome and a discrete choice set. Here, the choice $x_t \in \{0, 1\}$ is binary. Choosing $x_t = 1$ results in a gross random payoff θ_t and

[6] For example, endowing this model with a network structure, as in Gale and Kariv (2003), might allow for parameters that more readily map into observable data.

[7] It is worth mentioning that this guarantees that agents will not experiment. Experimentation will not yield any benefits, because an agent has no future period of action in which to use the information collected from experimenting. This particular result, and others, is explored the literature on optimal experimentation. See for instance, Aghion et al. (1991), or the material and references in the previous chapter. It may be interesting to study the effects of giving agents the incentive to experiment in this model. We leave that for future research.

cost c. Choosing $x_t = 0$ yields a payoff of 0. In other words,

$$u(x_t;\theta_t) = \begin{cases} \theta_t - c, & \text{if } x_t = 1 \\ 0, & \text{if } x_t = 0 \end{cases}.$$

Let $\theta_t = \gamma_t\phi$ so that the gross payoff is a composite of some time-varying object $\gamma_t \in \left\{\gamma^H, \gamma^L\right\}$, which we identify with institutions, and some underlying state of the world $\phi \in \left\{\phi^H, \phi^L\right\}$. We assume it is common knowledge that the γ_t's are independent and identically distributed[8] with $\Pr\left(\gamma_t = \gamma^H\right) = 1 - \varepsilon$.

We make the natural assumptions that $\phi^H > \phi^L$ and $\gamma^H > \gamma^L$, and that $\gamma^L\phi^L < c < \gamma^H\phi^H$, so that preferences are non-degenerate.

One might be inclined to think that institutions should be endogenous. There are a few reasons to abstract from that. First, even if institutions are endogenous, their impacts on specific individuals are still, in some sense, random, so it seems reasonable to model it as we do here. For example, even within the *mita* catchment, only one-seventh of the adult male population was forced to work in the Potosi and Huancavelica mines. Moreover, many of the institutions considered in the literature are colonial in origin. This means that to many of the people they affected, these institutions were genuinely exogenous. In section 3.3.2, we consider one method of endogenising γ_t.

If $\gamma^L\phi^H < c$ and $\gamma^H\phi^L < c$, then states of the world and institutions are complements. One needs both institutions and the state of the world to be favourable to achieve positive payoffs. On the other hand, setting $\gamma^L\phi^H > c$ and $\gamma^H\phi^L > c$ means that institutions and states of the world are substitutes. In this case, an individual needs only one of institutions or the state of the world to be favourable for $x_t = 1$ to have a higher payoff. If $\gamma^L\phi^H > c > \gamma^H\phi^L$, then we say that institutions are degenerate: they would have no impact on choices if there were no uncertainty. We remain agnostic about these assumptions regarding the interaction between states and institutions for the moment, but examine the impact of these different conditions later.

Intuitively, one might think of choosing $x_t = 1$ as the choice to exert high effort over one's lifetime, and the payoffs to exerting that effort depend on how one is affected by institutions (which change over time) and the (static) state of the world. We associate $\gamma_t = \gamma^L$ with being exposed to extractive institutions, and ε with the persistence of extractive institutions. Moreover, one decides based on information about the world derived from the observable institutions and an unobservable state of the world. One's information about the state of the world then comes from a noisy signal of the state, and

[8] It might be more natural to model γ_t as a Markov process with $\Pr\left(\gamma_{t+1} = \gamma_t | \gamma_t\right) = 1 - \varepsilon$, which better captures the notion that institutions are persistent. However, this assumption very quickly makes the model intractable. We consider this possibility with additional tractability assumptions in an extension in section 3.3.1.

the actions of one's predecessors (which we could think of as the individual's parents, grandparents, and so on).

At time t, individual t chooses x_t after observing a private signal $s_t \in \left\{s^L, s^H\right\}$; $h_t = \{x_1, \ldots, x_{t-1}\}$, the history of all preceding actions; and γ_t. Individual t does not observe $\gamma_{t'}$ or $s_{t'}$ for $t' \neq t$.[9]

Conditional on the state, all signals are iid:

$$\Pr\left(s_t = s^i | \phi = \phi^i\right) = \alpha > \frac{1}{2}$$

for $i \in \{H, L\}$. Define $p_t = \Pr\left(\phi = \phi^H | h_t\right)$, the public belief on the state in period t. Suppose as well that there is a common prior belief $p_1 = \Pr\left(\phi = \phi^H\right)$.

3.2.2 Individual Behaviour

Consider now the belief of individual t. Their belief on the state, upon receiving signal $s_t = s^i$, is

$$\begin{aligned} q_t^i &= \Pr\left(\phi = \phi^H | h_t, s_t = s^i\right) \\ &= \frac{\Pr\left(\phi = \phi^H, s_t = s^i | h_t\right)}{\Pr\left(s_t = s^i | h_t\right)} \\ &= \frac{\Pr\left(s_t = s^i | \phi = \phi^H, h_t\right) \Pr\left(\phi = \phi^H | h_t\right)}{\Pr\left(s_t = s^i | h_t\right)} \end{aligned}$$

where the second and third equalities follow from repeated application of Bayes' rule. Using the definitions given above, and the fact that signals are independent conditional on the state, we have that[10]

$$\begin{aligned} q_t^H &= \frac{\alpha p_t}{\alpha p_t + (1-\alpha)(1-p_t)} \\ q_t^L &= \frac{(1-\alpha) p_t}{(1-\alpha) p_t + \alpha (1-p_t)}. \end{aligned}$$

Each agent maximises

$$E\left[u(x_t; \theta_t) | s_t = s^i, \gamma_t = \gamma^j, h_t\right] = \begin{cases} \gamma^j \left[q_t^i \phi^h + \left(1 - q_t^i\right) \phi^L\right], & \text{if } x_t = 1 \\ 0, & \text{if } x_t = 0 \end{cases}.$$

[9] While overall historical institutions are typically close to perfectly observable, there tends to be more uncertainty over their impact on specific individuals. As such, the assumption of the non-observability of historical γ_t does not, as a first approximation, seem unreasonable.

[10] The history depends on the state only through the past signals. Hence, conditional on the state, the current signal is independent of history as well.

This yields a simple decision rule: given $\gamma_t = \gamma^j$, $x_t = 1$ only if $q_t^i > \Theta^j$, where[11]

$$\Theta^j = \frac{\frac{c}{\gamma^j} - \phi^L}{\phi^H - \phi^L}.$$

The previous inequality is a necessary condition on the private belief q_t^i for individual t to choose $x_t = 1$. The sufficient conditions for when she does choose $x_t = 1$ depend on assumptions about her actions when she is indifferent, which we now specify, along with her complete decision rule.

Suppose that the agent observes $\gamma_t = \gamma^j$ and receives signal $s_t = s^L$. This agent chooses $x_t = 1$ if and only if

$$q_t^L > \Theta^j \iff p_t > \frac{\alpha\Theta^j}{\alpha\Theta^j + (1-\alpha)\left(1-\Theta^j\right)} = \bar{p}^j.$$

Suppose instead that the signal received was $s_t = s^H$. The agent chooses $x_t = 1$ if and only if

$$q_t^H \geq \Theta^j \iff p_t \geq \frac{(1-\alpha)\,\Theta^j}{\alpha\left(1-\Theta^j\right) + (1-\alpha)\,\Theta^j} = \underline{p}^j.$$

We are assuming, without loss[12], that the action choice when indifferent minimises the possibility of herding. It is easy to see, given the assumptions $\alpha > 1/2$ and $\gamma^H > \gamma^L$ that $\bar{p}^j > \underline{p}^j$, $\bar{p}^L > \bar{p}^H$, and $\underline{p}^L > \underline{p}^H$. Moreover, as long as $\alpha$is not too large, we would have that $\underline{p}^L > \bar{p}^H$.[13] This means that the action in period t is partially informative of s_t for at least some $p_t \in \left[\underline{p}^H, \bar{p}^L\right]$. However, outside this region, actions are completely uninformative about signals. To see this, note that for $p_t < \bar{p}^H$, individual t will choose $x_t = 0$ regardless of her signal or observed institutions. Similarly, for $p_t > \bar{p}^L$, any agent will choose $x_t = 1$.

When institutions are degenerate, $0 < \underline{p}^H < \bar{p}^H < \underline{p}^L < \bar{p}^L < 1$. If institutions and the state of the world are complements, then $0 < \underline{p}^H < \bar{p}^H < 1 < \underline{p}^L < \bar{p}^L$. If institutions and states of the world are substitutes, $\underline{p}^H < \bar{p}^H < 0 < \underline{p}^L < \bar{p}^L < 1$. The succeeding discussion allows fro any of these three cases.

---

[11]This is similar to the analysis in Moscarini et al. (1998), except that $\Theta^j = 1/2$.
[12]See Moscarini et al. (1998).
[13]The exact condition is that

$$\alpha < \frac{1}{1+\sqrt{Z}},$$

where

$$Z = \frac{\Theta^H}{\Theta^L}\frac{1-\Theta^L}{1-\Theta^H} < 1.$$

We assume that this holds throughout.

3.2.3 Belief Dynamics

We adapt the definition of informational cascades in BHW to a form appropriate to our setting. Intuitively, an informational cascade occurs in some period when the action in that period cannot provide additional information about the state ϕ.

Definition 3.1. There is an *informational cascade* (or cascade) on action x at time t whenever, conditional on γ_t, action x is taken independently of individual t's private signal s_t.

A cascade on $x = 1$ occurs as soon as $p_t > \bar{p}^L$, whereas a cascade on $x = 0$ occurs as soon as $p_t < \underline{p}^H$. Moreover, careful consideration of the decision rule given above shows that, for $p_t \in \left(\bar{p}^H, \underline{p}^L\right)$, actions are dependent solely on γ_t in that $x_t = 1$ if and only if $\gamma_t = \gamma^H$. Hence, it makes sense to define the cascade set $X = \hat{X} \cap [0,1]$, where

$$\hat{X} = \left(-\infty, \underline{p}^H\right) \cup \left(\bar{p}^H, \underline{p}^L\right) \cup \left(\bar{p}^L, \infty\right).$$

On X, actions are uninformative about signals, so the public belief must be stationary: $p_{t+1} = p_t$. Outside the cascade set X, actions are partially informative about signals. Whenever $p_t \notin X$, observing $x_t = 1$ can reveal information about γ_t and s_t, depending on the value of p_t. For instance, if $p_t \in \left[\underline{p}^H, \bar{p}^H\right]$, an individual chooses $x_t = 1$ if and only if $\gamma_t = \gamma^H$ and $s_t = s^H$. One can reason similarly about $\left[\underline{p}^L, \bar{p}^L\right]$. The public belief is therefore updated according to the following formula:

If $x_t = 1$,

$$p_{t+1} = \begin{cases} q_t^H, & \text{if } p_t \in \left[\underline{p}^H, \bar{p}^H\right] \\ \frac{\left[(1-\varepsilon)+\alpha\varepsilon\right]p_t}{(1-\varepsilon)+\varepsilon\left[\alpha p_t+(1-\alpha)(1-p_t)\right]}, & \text{if } p_t \in \left[\underline{p}^L, \bar{p}^L\right] \end{cases}.$$

If $x_t = 0$,

$$p_{t+1} = \begin{cases} \frac{\left[\varepsilon+(1-\varepsilon)(1-\alpha)\right]p_t}{\varepsilon+(1-\varepsilon)\left[\alpha(1-p_t)+(1-\alpha)p_t\right]}, & \text{if } p_t \in \left[\underline{p}^H, \bar{p}^H\right] \\ q_t^L & \text{if } p_t \in \left[\underline{p}^L, \bar{p}^L\right] \end{cases}.$$

This is easily derived using Bayes' rule. We summarise this result below.

Proposition 3.1. *The public belief p_t is updated in every period according to*

$$p_{t+1} = \begin{cases} q_t^H, & \text{if } x_t = 1 \text{ and } p_t \in \left[\underline{p}^H, \bar{p}^H\right] \\ \frac{\left[\varepsilon+(1-\varepsilon)(1-\alpha)\right]p_t}{\varepsilon+(1-\varepsilon)\left[\alpha(1-p_t)+(1-\alpha)p_t\right]} & \text{if } x_t = 0 \text{ and } p_t \in \left[\underline{p}^H, \bar{p}^H\right] \\ \frac{\left[(1-\varepsilon)+\alpha\varepsilon\right]p_t}{(1-\varepsilon)+\varepsilon\left[\alpha p_t+(1-\alpha)(1-p_t)\right]} & \text{if } x_t = 1 \text{ and } p_t \in \left[\underline{p}^L, \bar{p}^L\right] \\ q_t^L & \text{if } x_t = 0 \text{ and } p_t \in \left[\underline{p}^L, \bar{p}^L\right] \\ p_t & \text{if } p_t \in X \end{cases},$$

where X is the cascade set defined earlier.

We are now ready to state our main result, which is similar to that found in the literature on social learning.

Proposition 3.2. *p_t converges almost-surely to some limiting random variable p_∞ that lies (almost-surely) in the cascade set X. Thus, cascades occur with probability one.*

Proof. p_t is bounded by definition, and a martingale by construction. The convergence result then follows from the martingale convergence theorem. The second part of the proposition follows from the fact that outside X, $|p_{t+1} - p_t|$ is bounded below by some positive constant. Convergence therefore is possible only if the public belief enters the cascade set X. □

There are alternative proofs of belief convergence implying action convergence in Bikhchandani et al. (1992); Smith et al. (2017); Smith and Sørensen (2000).

The previous discussion makes it immediately apparent why we are only interested in the cases of complementarity, substitutability, and degeneracy. Suppose, on the contrary, that $\gamma^L \phi^H < c < \gamma^H \phi^L$. One should expect this specification means that the state of the world doesn't matter. Indeed, under this assumption, we would have that $\Theta^L > 1$ and $\Theta^H < 0$, which then means that $\bar{p}^H < 0$ and $\underline{p}^L > 1$. The unit interval is therefore a strict subset of the cascade set, and we would have immediate belief convergence to the prior p_1. This case is clearly uninteresting, so we focus on other cases from now.

3.2.4 Pessimistic Herds and Luck

We now state a result on what types of informational cascades are possible given beliefs p_t. Recall that $\underline{p}^H < \bar{p}^H < \underline{p}^L < \bar{p}^L$.

Proposition 3.3. *Suppose that $\alpha < 1/\left(1 + \sqrt[3]{Z}\right)$, where $Z = \frac{\Theta^H}{\Theta^L}\frac{1-\Theta^L}{1-\Theta^H}$. For $p_t \in \left[0, \underline{p}^L\right)$, we have that $p_\infty \notin \left(\bar{p}^L, 1\right]$. Similarly, for $p_t \in \left(\bar{p}^H, 1\right]$, $p_\infty \notin \left[0, \underline{p}^H\right)$.*

Proof. If $p_t \in X$, then $p_\infty = p_t$. Suppose $p_t \notin X$. We consider the case when $p_t \in \left[\underline{p}^L, \bar{p}^L\right]$. The reasoning for when $p_t \in \left[\underline{p}^H, \bar{p}^H\right]$ is similar.

The result follows from the fact that no public belief $p_t \in \left[\underline{p}^L, \bar{p}^L\right]$ can jump over the interval $\left(\bar{p}^H, \underline{p}^L\right)$ after observing $x_t = 0$. To see why, notice that, even $p_t = \underline{p}^L$, $p_{t+1} > \bar{p}^H$. This completes the proof. □

Definition 3.2. A *herd* takes place at time T if all actions after time T are identical: for $t > T$, $x_t = x_T$. A herd on $x = 0$ is a *pessimistic herd*, and a herd on $x = 1$ is an *optimistic herd*.

A pessimistic herd occurs as soon as $p_t \in \left[0, \underline{p}^H\right)$. Given proposition 3.3, it is easy to see that, for a low enough prior[14], and when institutions and states are not substitutes, pessimistic herds happen with strictly positive probability, even if the state is ϕ^H. In particular, a long enough sequence of bad draws of γ_t will lead to a pessimistic herd. Crucially, this holds regardless of the actual draws of signals s_t. To see this, recognise that given the decision rule derived above, whenever $\gamma_t = \gamma^L$, $x_t = 0$, as long as $p_1 \in \left[\underline{p}^H, \bar{p}^H\right]$ (i.e the prior was in the region that admits a pessimistic herd). Hence, any period where $\gamma_t = \gamma^L$ will lead to downward adjustment of beliefs.

This plausibly sheds light on a mechanism that would explain why institutions often have effects that persist long after the institutions associated with those effects have been changed or reformed. Even if institutions are reformed so that $\varepsilon = 0$, there will be no effect on a pessimistic herd. This is because the cascade set is unaffected by changes in ε.

Given this result, we can now interpret the patterns in the data as resembling a pessimistic herd. One can think of the exposure to extractive institutions such as the *mita* or the slave trade s a long sequence of bad draws of γ_t which then leads to a cascade that entails herding on $x_t = 0$. The herding on low effort is then what causes bad outcomes for the victims of extractive institutions.

This model also suggests that luck is also an important factor in determining the long-term impact of institutions. To be specific, there has to be an element of bad luck involved in creating a pessimistic herd, since there needs to be a sequence of bad draws of signals or institutions. This is contrary to the argument in Acemoglu and Robinson (2012), where they claim that institutions are much more important than luck in determining long-term economic outcomes.[15]

3.2.5 Complementarity and Substitutability

Pessimistic herds can only happen under a low initial prior p_1. In particular, we must have that $p_1 \in \left[0, \bar{p}^H\right]$ in order for a pessimistic herd to occur. However, we can also derive parameter restrictions for which the set of priors that make a pessimistic herd possible is comparatively large, and hence, in some sense, more likely. In particular, if we assume that good institutions and favourable states of the world are complements, so that $\gamma^L \phi^H < c$ and $\gamma^H \phi^L < c$, then $\underline{p}^L > 1$ and $\bar{p}^H$ is close to 1. Moreover, under this complementarity assumption, then optimistic herds are impossible, as given

[14] We must have that $p_1 \in \left[0, \bar{p}^H\right]$. Otherwise there will never be a pessimistic herd.

[15] Even if one is not persuaded by the mechanism here, luck is almost certainly important in the selection of institutions. Hence, luck may still be much more important than institutions. However, given that institutions themselves are often thought of as equilibrium outcomes (Greif and Kingston, 2011), and luck can be thought of in terms of selection of equilibria, it was never clear what it meant for institutions to be more important than luck.

by proposition 3.3. One can also see this by considering belief dynamics in this case, with[16] $X^C = \left[0, \underline{p}^H\right) \cup \left(\bar{p}^H, 1\right]$:

$$p_{t+1} = \begin{cases} q_t^H & \text{if } x_t = 1 \text{ and } p \notin X^C \\ \frac{\left[\varepsilon + (1-\varepsilon)(1-\alpha)\right] p_t}{\varepsilon + (1-\varepsilon)\left[\alpha(1-p_t) + (1-\alpha)p_t\right]} & \text{if } x_t = 0 \text{ and } p_t \notin X^C \\ p_t, & \text{if } p_t \in X^C \end{cases}.$$

Here, when beliefs enter $p_t \in \left(\bar{p}^H, 1\right]$ there is an informational cascade, but agents do not herd on any specific action. Individuals choose the action based solely on their observed realisation of γ_t.

Hence, when institutions and states are complements, pessimistic herds are more likely to occur.

Suppose instead that $\gamma^L \phi^H > c$ and $\gamma^H \phi^L > c$, so that institutions and states of the world are substitutes. Public beliefs now evolve according to

$$p_{t+1} = \begin{cases} \frac{\left[(1-\varepsilon) + \alpha\varepsilon\right] p_t}{(1-\varepsilon) + \varepsilon\left[\alpha p_t + (1-\alpha)(1-p_t)\right]} & \text{if } x_t = 1 \text{ and } p_t \notin X^S \\ q_t^L & \text{if } x_t = 0 \text{ and } p_t \notin X^S \\ p_t & \text{if } p_t \in X^S \end{cases},$$

where[17] $X^S = \left[0, \underline{p}^L\right) \cup \left(\bar{p}^L, 1\right]$.

In this environment, pessimistic herds are now impossible. However, there are some interesting results in this set up that we now describe. Instead of considering pessimistic herds, which do not occur, we focus on pessimistic belief cascades, which can still occur. This happens when beliefs enter $\left[0, \bar{p}^L\right)$.

Notice that we can recover the canonical social learning model by setting $\varepsilon = 1$.This observation highlights the impact of the way we incorporate institutions here to the basic model. Suppose $p_t \notin X^S$ so that learning still occurs. If $\varepsilon < 1$, with positive probability, one does not become victim to extractive institutions. Relative to the case when $\varepsilon = 1$, the increase in p_t after observing $x_t = 1$ is slowed. The decrease in p_t after observing $x_t = 0$, however, remains the same. Formally, treating the public belief as a function of ε, $p_t(\varepsilon) \leq p_t(1)$ for any fixed history h_t. This means, somewhat counterintuitively, that pessimistic cascades are more likely when there is some possibility that one is exposed to non-extractive institutions rather than when institutions are always extractive.

[16] Recall that, under complementarity, $0 < \underline{p}^H < \bar{p}^H < 1 < \underline{p}^L$.
[17] Recall that when institutions and states are substitutes, $\bar{p}^H < 0 < \underline{p}^L < \bar{p}^L < 1$.

Moreover, in any time period t, we can define the speed of learning on ϕ^H as

$$v^H = \frac{1}{t^+} \sum_{k=1}^{t} \max\{p_{k+1} - p_k, 0\}$$

where $t^+ = \#\{k \leq t : p_{k+1} > p_k\}$.[18]

v^H is smallest when ε is very close to 0. This means that pessimistic beliefs are most likely when the probability of being exposed to extractive institutions is very small, or when extractive institutions are not very persistent. In light of the data, one might view this as suggestive of good institutions not being substitutes for a favourable state of the world. Another way of thinking about this result is that it is an explanation of why certain types of institutional reform improve economic outcomes while others don't. The reforms that fail represent institutional reforms that are complementary to a good state of the world, when the reform that is needed is the creation of institutions that are substitutes for a favourable state.

Notice, however, that actions in this model are driven purely by beliefs and not outcomes. Therefore, if people systematically underestimate ε, and there is empirical evidence that they might, then given this model, people possessing mistakenly pessimistic beliefs about the world does not seem unlikely.[19] This result can then be incorporated into a richer model that can produce a pessimistic cascade on actions, instead of just beliefs. We leave that idea for future research.

3.3 Extensions

3.3.1 Institutions as a Markov Process

Suppose that, instead of being iid, institutions followed a Markov process, so that $\Pr\left(\gamma_{t+1} = \gamma_t | \gamma_t\right) = 1 - \varepsilon$. One might view this as more natural than our initial assumption because it better captures the notion that institutions are persistent. However, this assumption creates significant difficulty in characterising public belief dynamics, which will now no longer follow a Markov process, and instead are fully dependent on complete histories. However, making additional assumptions recovers tractability. Instead of conditioning the public belief solely on the history, define it to be $p_t = \Pr\left(\phi = \phi^H | h_t, \gamma_t\right)$. It is reasonably easy to see that public beliefs are once again Markov, and hence can be characterised in a fashion that is qualitatively similar to the analysis in the previous section

It is worth pointing out that this is akin to a bounded rationality assumption: this assumption entails that the public is forgetful. In particular, it means that the public

[18]Think of this as the average increment in p_t every time beliefs increase.

[19]The psychology literature documents this as the just-world fallacy (Montada and Lerner, 1998). Bénabou and Tirole (2006) apply this finding to an economic model.

observes the effect of institutions in a period t, and uses that information to update their beliefs on the state of the world. However, in period $t+1$, the public no longer has the information about institutions in period t to update their beliefs with.

The forgetfulness assumption is important. If the public did not forget the values of $\gamma_{t'}$ for $t\prime < t$, then the signal-jamming effect of institutions disappears, and the results derived in the previous section that depart from the standard social learning model disappear.

3.3.2 Investment in Institutions

n this section, we sketch a variant of the model where γ_t is endogenous. Suppose now that the effect of institutions γ_t is now a deterministic function of investment $i_t \in 0, 1$. In particular, we have that

$$\gamma_t = \begin{cases} \gamma^H, & \text{if } i_t = 1 \\ \gamma^L, & \text{if } i_t = 0 \end{cases}.$$

However, agents also face an unknown cost of investment, $\hat{c} \in \left\{\hat{c}^0, \hat{c}^1\right\}$, with $\gamma^L \varphi^L < \hat{c}^0 < \hat{c}^1$ and $\hat{c}^0 < \gamma^H \phi^H$. Let there be a common prior on $\hat{c}$ given by $r_1 = \Pr\left(\hat{c} = \hat{c}^0\right)$, and assume for simplicity that costs are independent of ϕ. Preferences are now given by

$$u\left(x_i, i_t\right) = \begin{cases} \left[\left(\gamma^H - \gamma^L\right)\phi - \hat{c}\right] i_t + \gamma^L \phi - c, & \text{if } x_t = 1 \\ -\hat{c} i_t, & \text{if } x_t = 0 \end{cases}.$$

Each agent receives two signals, s_t^ϕ and s_t^c. s_t^ϕ is the same as s_t in previous sections, and $s_t^c \in \left\{s^0, s^1\right\}$ is now a signal about $\hat{c}$. In particular, we assume that $\Pr\left(s_t^c = s^j | \hat{c} = \hat{c}^j\right) = \beta > \frac{1}{2}$. Individual t observes $h_t = \{i_1, x_1, \ldots, i_{t-1}, x_{t-1}\}$. This gives the public beliefs $p_t = \Pr\left(\phi = \phi^H | h_t\right)$ and $r_t = \Pr\left(\hat{c} = \hat{c}^0 | h_t\right)$.

The timing of the model in period t is as follows. Individual t first receives the signal s_t^c, which she uses to update $u_t^j = \Pr\left(\hat{c} = \hat{c}^0 | h_t, s_t^c = s^j\right)$, her private belief about $\hat{c}$. Then, after choosing i_t, she receives the signal s_t^ϕ, which is used to update her private belief $q_t^j = \Pr\left(\phi = \phi^H | h_t, s_t^\phi = s^j\right)$, and chooses x_t.

As before, we have that

$$\begin{aligned} u_t^0 =& \frac{\beta r_t}{\beta r_t + (1-\beta)(1-r_t)} \\ u_t^1 =& \frac{(1-\beta) r_t}{\beta(1-r_t) + (1-\beta) r_t} \\ q_t^H =& \frac{\alpha p_t}{\alpha p_t + (1-\alpha)(1-p_t)} \\ q_t^L =& \frac{(1-\alpha) p_t}{\alpha(1-p_t) + (1-\alpha) p_t}. \end{aligned}$$

The decision rule for x_t is very similar to what we had earlier. Suppose first that $i_t = 1$. Given $s_t^\phi = s^H$, individual t chooses $x_t = 1$ if and only if $q_t^H \geq \Theta^H$. Given s_t^L, individual t chooses $x_t = 0$ if and only if $q_t^L > \Theta^H$. Suppose instead that $i_t = 0$. Given $s_t^\phi = s^H$, $x_t = 1$ is chosen if and only if $q_t^H \geq \Theta^L$. When $s_t^\phi = s^L$, $x_t = 1$ if and only if $q_t^L > \Theta^L$.

The decision rule for i_t is less obvious. Consider first the expected payoff from choosing $i_t = 1$:

$$\max\left\{\gamma^H E\left[\phi|h_t\right] - c, 0\right\} - E\left[\hat{c}|h_t, s_t^c\right].$$

Contrast this with the expected payoff from $i_t = 0$:

$$\max\left\{\gamma^L E\left[\phi|h_t\right] - c, 0\right\}.$$

Thus, $i_t = 1$ only if

$$\gamma^H E\left[\phi|h_t\right] - E\left[\hat{c}|h_t, s_t^c\right] > \max\left\{\gamma^L E\left[\phi|h_t\right], c\right\}.$$

Otherwise, the agent will at most be indifferent between investing and not investing. We can proceed in the same way as earlier to describe t's decision rule for i_t. Define

$$\hat{C}\left(h_t\right) = \frac{\hat{c}^1 - \left[\gamma^H E\left[\phi|h_t\right] - \max\left\{\gamma^L E\left[\phi|h_t\right], c\right\}\right]}{\hat{c}^1 - \hat{c}^0}.$$

Suppose $s_t^c = s^0$. Then, she invests if and only if

$$u_t^0 \geq \hat{C}\left(h_t\right) \iff r_t \geq \frac{(1-\beta)\hat{C}\left(h_t\right)}{\beta\left[1 - \hat{C}\left(h_t\right)\right] + (1-\beta)\hat{C}\left(h_t\right)} = \underline{r}\left(h_t\right).$$

If instead $s_t^c = s^1$, she invests if and only if

$$u_t^1 > \hat{C}\left(h_t\right) \iff r_t > \frac{\beta\hat{C}\left(h_t\right)}{\beta\hat{C}\left(h_t\right) + (1-\beta)\left[1 - \hat{C}\left(h_t\right)\right]} = \bar{r}\left(h_t\right).$$

The dependence of the belief thresholds $\bar{r}$ and $\underline{r}$ on h_t make fully characterising belief dynamics and the cascade set quite difficult. However, we can show that, under certain conditions, the cascade set is non-empty, which means that we can still have (non-trivial) information cascades in this environment.

Assume that

$$\hat{c}^1 > \gamma^H\phi^H - \max\left\{\gamma^L\phi^H, c\right\}$$

and that

$$\gamma^H\phi^L - c > \hat{c}^0.$$

This guarantees that for any fixed sequence of actions h_t, $1-\delta > \bar{r}\left(h_t\right) > \underline{r}\left(h_t\right) > \delta$

for some $\delta > 0$. This means that

$$\bar{R} = \limsup_{t\to\infty} \bar{r}(h_t) < 1$$
$$\underline{R} = \liminf_{t\to\infty} \underline{r}(h_t) > 0.$$

Thus, $[0, \underline{R}) \cup (\bar{R}, 1]$ is a subset of the cascade set. A long enough sequence of bad signals s_t^ϕ and s_t^c clearly leads to an informational cascade with herding on $x_t = 0$ and $i_t = 0$. In other words, if an early sequence of agents is sufficiently pessimistic, there will be a herd of no investment and low effort.

3.4 Discussion

3.4.1 Robustness of Wrong Cascades

The results we derive are clearly heavily reliant on BHW's result of the possibility of incorrect cascades. However, this result is heavily reliant on the coarseness of both the action space (Lee, 1993) and signal space (Smith and Sørensen, 2000). In particular, when there is a continuous action space or unbounded signals, beliefs will always converge to the truth. Hence, the assumptions made earlier are crucial to the results derived. This is certainly inconvenient, but not damning of the plausibility of social learning as a mechanism that drives the persistent effects of institutions.

First, it is not at all clear that the world and, particularly, individual perceptions of it are substantially richer than the discrete action-signal structure that we assume here. The fact that everyday language tends to be imprecise, for example, suggests that people are not accustomed to thinking and communicating with the precision that a continuous space requires. Indeed, it seems more likely that people think in terms of discrete objects and quantities rather than continuous ones. As such, modelling assumptions involving discrete objects are probably more accurate representations of reality than those that involve continuous ones, which are only typically assumed for tractability. As such, the canonical social learning model's lack of robustness to assumptions of discreteness is in no way an indictment of the model, or of results based on it that we derive.

Second, even in a model with a continuous action space and unbounded signals, learning about the true state of the world can be very slow when there is noise in the observation of other agents' signals. As such, from the point of view of welfare, even slow convergence to the truth is equivalent to convergence to a mistake. Moreover, slow learning about unknown parameters would be useless if those parameters change often enough, as they might in a richer model.[20] Vives (1993) demonstrates this in an economy with many competitive firms learning about some unknown parameter of the aggregate demand function through repeated market interaction.

[20]See, for example, Moscarini et al. (1998).

We now sketch how one might adapt the basic environment we have to accommodate a continuous action space with unbounded signals. Suppose that agents choose their effort level $x_t \in \mathbb{R}_+$, where their preferences are given by

$$u(x_t; \theta_t) = \theta_t x_t - c(x_t).$$

$c(x_t)$ is the (differentiable and strictly convex) cost of exerting effort x_t. Assume that $\theta_t = \gamma_t + \phi$, and that γ_t and ϕ are independent and normally distributed. Suppose that signals are given by $s_t = \phi + \eta_t$, where $\eta \sim N\left(0, \sigma_\eta^2\right)$, so that, conditional on ϕ, signals are normally distributed and independent across time. All other aspects of the basic model remain the same. The agent in every period then sets

$$x_t^* = \left(c'\right)^{-1}\left(\gamma_t + E\left[\phi | h_t, s_t\right]\right).$$

Standard results give us that

$$E\left[\phi | h_t s_t\right] = \alpha_t s_t + (1 - \alpha_t) E\left[\phi | h_t\right]$$

where

$$\alpha_t = \frac{\sigma_t^2}{\sigma_t^2 + \sigma_\eta^2},$$

and σ_t^2 is the variance of the public belief. The derivative of the cost function is uniquely invertible, so observing x_t^* is equivalent to observing $\gamma_t + \alpha_t s_t + (1 - \alpha_t) E\left[\phi | h_t\right]$, which is a noisy signal of s_t. Vives' results about slow convergence then apply.

3.4.2 Empirical Evidence

We now examine the literature more closely and consider the plausibility of the theory given the specific empirical findings in Nunn (2008); Nunn and Wantchekon (2011); Dell (2010). Nunn (2008); Nunn and Wantchekon (2011) typify the research that presents a mechanism that is reminiscent of the theory we develop, whereas Dell (2010) gives an example of a mechanism that is not.

The mechanism proposed here is closest to that suggested by Nunn and Wantchekon (2011). They argue that one channel by which the slave trade caused depressed economic outcomes in Africa today is mistrust. They find that current differences in trust levels within Africa can be traced back to the transatlantic and Indian Ocean slave trades. Today, members of ethnic groups that were most exposed to the slave trade are also the least trusting of their relatives, neighbours, co-ethnics and local governments. This complements Nunn (2008), where he shows that the slave trade appears to have had a causal impact on African economic development.

One can therefore think of the state of the world ϕ as the inherent level of trustworthiness of other people, and γ_t as the level of one's exposure to the slave trade. Hence, a long enough sequence of heavy exposure to the slave trade, as shown earlier, will lead to pessimistic beliefs about how trustworthy other people are. As these beliefs grow ever more pessimistic, they eventually enter the cascade set where beliefs become invariant to institutional reform. We would then have herding on the low-payoff outcome, regardless of how inherently trustworthy other people actually are. One can thus think of choosing $x_t = 1$ as the choice to cooperate in some stage cooperation game. We do not model this cooperation game, but it is easy to see how our framework can be extended to this environment. In fact, if we assume that individuals only observe the actions of one's predecessors, and not their predecessors' opponents in each stage game, this extension is trivial.

Moreover, the mechanism Nunn and Wantchekon suggest involves a hallmark of boundedly rational agents, heuristic "rules-of-thumb" of what turns out to be social learning. The analysis in the previous sections shows, however, that bounded rationality is not necessary for behavioural persistence.[21]

On the other hand, Dell's findings about channels of persistence of the *mita* appear to not be in line with the theory proposed here. Dell claims that the *mita* prevented the emergence of large landowners within the *mita* catchment. The absence of large landowners meant that there was no one who could lobby for or provide public goods, such as roads. The lack of these public goods today prevents participation in agricultural markets. Households outside the *mita* catchment, on the other hand, are able to participate in said agricultural markets, and hence enjoy much better economic outcomes.

Even though this mechanism does not line up cleanly with the mechanism presented in the model, it does not quite rule out social learning as an explanation of persistence. In particular, it is possible that social learning effects only play a second-order role in the case of the *mita*. It is also possible that a social learning effect is lurking in the background: Dell does not provide an explanation as to what the channels through which under-provision of public goods persists. In particular, social learning might explain why there is persistent under-provision of public goods within the *mita* catchment area, as in the model in section3.3.2. In particular, one can think of x_t as the decision to participate in agricultural markets, whereas i_t represents investment in public goods provision.

3.5 Conclusion

This paper proposes a theory of the persistent effects of institutions via social learning. To do this, we adapted the model by BHW to a setting where institutions affect both

[21] Strictly speaking, the model Nunn and Wantchekon suggest as an underlying mechanism (Boyd and Richerson, 1995) does not make bounded rationality assumptions, but rather uses methods from evolutionary game theory. The analysis we conducted earlier shows that the mechanism can also be driven by standard methods of modelling individual behaviour.

individual decisions and outcomes. We showed that the persistent underdevelopment that traces back to extractive historical institutions can be thought of as herding on a bad outcome. This herding is driven by an informational cascade of pessimistic beliefs about the state of the world. A long enough sequence of agents exposed to extractive institutions will create a pessimistic informational cascade. Moreover, once agents start herding, reforms of institutions will not have any impact on the herd.

A cascade of pessimistic beliefs is possible only if institutions are degenerate, or when institutions are only complementary to a favourable state of the world. Institutions are degenerate when they would have no impact on individual behaviour if not for uncertainty about the state of the world. In both cases, the prior on the state of the world must be sufficiently low, but less so under complementarity than under degeneracy. In this sense, pessimistic herding is most likely when institutions are complementary to the state of the world.

The model appears to be consistent with the empirical evidence on the effects of historical institutions. A more systematic analysis of how the model fares with respect to the data would be interesting, but this is left for future research.

Bibliography

Abreu, D. (1986). Extremal equilibria of oligopolistic supergames. *Journal of Economic Theory*, 39(1):191 – 225.

Abreu, D., Milgrom, P., and Pearce, D. (1991). Information and timing in repeated partnerships. *Econometrica*, 59(6):1713–1733.

Abreu, D., Pearce, D., and Stacchetti, E. (1986). Optimal cartel equilibria with imperfect monitoring. *Journal of Economic Theory*, 39(1):251 – 269.

Acemoglu, D. (2003). Why not a political Coase theorem? social conflict, commitment, and politics. *Journal of Comparative Economics*, 31(4):620–652.

Acemoglu, D. and Jackson, M. (2015). History, expectations, and leadership in the evolution of social norms. *The Review of Economic Studies*, 82(2).

Acemoglu, D. and Johnson, S. (2005). Unbundling institutions. *Journal of Political Economy*, 113(5):949–995.

Acemoglu, D., Johnson, S., and Robinson, J. A. (2001). The colonial origins of comparative development: An empirical investigation. *American Economic Review*, 91(5):1369–1401.

Acemoglu, D., Johnson, S., and Robinson, J. A. (2002). Reversal of fortune: Geography and institutions in the making of the modern world income distribution. *The Quarterly Journal of Economics*, 117(4):1231–1294.

Acemoglu, D. and Robinson, J. A. (2001). A theory of political transitions. *American Economic Review*, 91(4):938–963.

Acemoglu, D. and Robinson, J. A. (2008). Persistence of power, elites, and institutions. *American Economic Review*, 98(1):267–293.

Acemoglu, D. and Robinson, J. A. (2012). *Why nations fail : the origins of power, prosperity, and poverty.* Profile, London.

Aghion, P., Bolton, P., Harris, C., and Jullien, B. (1991). Optimal learning by experimentation. *The Review of Economic Studies*, 58(4):621–654.

Argenziano, R., Severinov, S., and Squintani, F. (2016). Strategic information acquisition and transmission. *American Economic Journal: Microeconomics*, 8(3):119–55.

Aumann, R. J. (1964). Mixed and behavior strategies in infinite extensive games. In Dresher, M. and Shapley, L. S., editors, *Advances in game theory*, pages 443–476. Princeton University Press.

Ball, I. (2017). Dynamic influence: persuasion and incentives. Working Paper.

Bandiera, O. and Rasul, I. (2006). Social networks and technology adoption in Northern Mozambique. *Economic Journal*, 116(514):869–902.

Banerjee, A. (1992). A simple model of herd behavior. *The Quarterly Journal of Economics*, 108(3).

Banerjee, A. and Iyer, L. (2005). History, institutions, and economic performance: The legacy of colonial land tenure systems in India. *American Economic Review*, 95(4):1190–1213.

Benabou, R. and Ok, E. A. (2001). Social mobility and the demand for redistribution: The Poum hypothesis. *The Quarterly Journal of Economics*, 116(2):447–487.

Bénabou, R. and Tirole, J. (2006). Belief in a just world and redistributive politics. *The Quarterly Journal of Economics*, 121(2):699–746.

Bergemann, D. and Bonatti, A. (2015). Selling cookies. *American Economic Journal: Microeconomics*, 7(3):259–94.

Bergemann, D., Brooks, B., and Morris, S. (2015). The limits of price discrimination. *American Economic Review*, 105(3):921–57.

Bergemann, D. and Hege, U. (2005). The financing of innovation: learning and stopping. *The RAND Journal of Economics*, 36(4):719–752.

Bergemann, D. and Morris, S. (2016a). Bayes correlated equilibrium and the comparison of information structures in games. *Theoretical Economics*, 11(2):487–522.

Bergemann, D. and Morris, S. (2016b). Information design, Bayesian persuasion, and Bayes correlated equilibrium. *American Economic Review*, 106(5):586–91.

Bergemann, D. and Morris, S. (2017). Information design: a unifed perspective. Cowles Foundation Discussion Paper No. 2075R.

Bergemann, D. and Pesendorfer, M. (2007). Information structures in optimal auctions. *Journal of Economic Theory*, 137(1):580 – 609.

Bergemann, D. and Välimäki, J. (2002). Information acquisition and efficient mechanism design. *Econometrica*, 70(3):1007–1033.

Bergemann, D. and Välimäki, J. (2006). Information in mechanism design. In Blundell, R., Newey, W. K., and Persson, T., editors, *Advances in Economics and Econometrics: Theory and Applications, Ninth World Congress*, volume II, chapter 5, pages 186–221. Cambridge University Press.

Bergemann, D. and Välimäki, J. (2008). bandit problems. In Durlauf, S. N. and Blume, L. E., editors, *The New Palgrave Dictionary of Economics*. Palgrave Macmillan, Basingstoke.

Berry, D. A. and Fristedt, B. (1985). *Bandit problems: sequential allocation of experiments*. Monographs on statistics and applied probability. Springer.

Bikhchandani, S., Hirshleifer, D., and Welch, I. (1992). A theory of fads, fashion, custom, and cultural change as informational cascades. *Journal of Political Economy*, 100(5):992–1026.

Blackwell, D. (1951). Comparison of experiments. In *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, pages 93–102, Berkeley, Calif. University of California Press.

Bolton, P. and Dewatripont, M. (2005). *Contract Theory*. MIT Press.

Bolton, P. and Harris, C. (1999). Strategic experimentation. *Econometrica*, 67(2):349–374.

Bonatti, A. and Hörner, J. (2011). Collaborating. *The American Economic Review*, 101(2):632–663.

Bonatti, A. and Hörner, J. (2017a). Career concerns with exponential learning. *Theoretical Economics*, 12(1):425–475.

Bonatti, A. and Hörner, J. (2017b). Learning to disagree in a game of experimentation. *Journal of Economic Theory*, 169:234 – 269.

Boyd, R. and Richerson, P. J. (1995). Why does culture increase human adaptability? *Ethology and Sociobiology*, 16(2):125–143.

Chamley, C. (2004). *Rational Herds: Economic Models of Social Learning*. Cambridge University Press, Cambridge.

Che, Y.-K. and Hörner, J. (2017). Recommender systems as incentives for social learning. Yale University Working Paper.

Cialdini, R. B. and Goldstein, N. J. (2004). Social influence: Compliance and conformity. 55(1):591–621.

Clarke, F. H. (1990). *Optimization and nonsmooth analysis*, volume 5 of *Classics in applied mathematics*. Wiley, New York.

Coate, S. and Morris, S. (1999). Policy persistence. *American Economic Review*, 89(5):1327–1336.

Conley, T. G. and Udry, C. R. (2010). Learning about a new technology: Pineapple in Ghana. *American Economic Review*, 100(1):35–69.

Crawford, V. P. and Sobel, J. (1982). Strategic information transmission. *Econometrica*, 50(6):1431–1451.

Crémer, J. and McLean, R. P. (1985). Optimal selling strategies under uncertainty for a discriminating monopolist when demands are interdependent. *Econometrica*, 53(2):345–361.

Crémer, J. and McLean, R. P. (1988). Full extraction of the surplus in Bayesian and dominant strategy auctions. *Econometrica*, 56(6):1247–1257.

Cripps, M. and Thomas, C. (2016). Strategic experimentation in queues. Working Paper.

Cronshaw, M. B. and Luenberger, D. G. (1994). Strongly symmetric subgame perfect equilibria in infinitely repeated games with perfect monitoring and discounting. *Games and Economic Behavior*, 6(2):220 – 237.

Davis, M. (1993). *Markov models and optimization*. Monographs on statistics and applied probability. Springer-Science+Business Media.

Dell, M. (2010). The persistent effects of Peru's mining mita. *Econometrica*, 78(6):1863–1903.

Deutsch, M. and Gerard, H. B. (1955). A study of normative and informational social influences upon individual judgment. *The Journal of Abnormal and Social Psychology*, 51(3):629–636.

Di Pei, H. (2016). Reputation with strategic information disclosure. *Available at SSRN*.

Dong, M. (2016). Strategic experimentation with asymmetric information. Working Paper.

Dworczak, P. (2017). Mechanism design with aftermarkets: Cutoff mechanisms. Working paper, Northwestern University.

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

Eső, P. and Szentes, B. (2007). Optimal information disclosure in auctions and the handicap auction. *The Review of Economic Studies*, 74(3):705–731.

Evans, L. C. (2010). *Partial differential equations*, volume 19 of *Graduate studies in mathematics*. American Mathematical Society, 2nd edition.

Fernandez, R. and Rodrik, D. (1991). Resistance to reform: Status quo bias in the presence of individual-specific uncertainty. *The American Economic Review*, 81(5):1146–1155.

Field, E. (2007). Entitled to work: Urban property rights and labor supply in peru. *The Quarterly Journal of Economics*, 122(4):1561–1602.

Forges, F. (1986). An approach to communication equilibria. *Econometrica*, 54(6):1375–1385.

Fudenberg, D. and Yamamoto, Y. (2010). Repeated games where the payoffs and monitoring structure are unknown. *Econometrica*, 78(5):1673–1710.

Fudenberg, D. and Yamamoto, Y. (2011). Learning from private information in noisy repeated games. *Journal of Economic Theory*, 146(5):1733 – 1769.

Gale, D. and Kariv, S. (2003). Bayesian learning in social networks. *Games and Economic Behavior*, 45(2):329–346.

Galor, O., Moav, O., and Vollrath, D. (2009). Inequality in landownership, the emergence of human-capital promoting institutions, and the great divergence. *Review of Economic Studies*, 76(1):143–179.

Georgiadis, G. and Szentes, B. (2018). Optimal monitoring design. Working paper.

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

Gittins, J., Glazebrook, K., and Weber, R. (2011). *Multi-armed bandit allocation indices*. Wiley-Blackwell, 2nd edition.

Glaeser, E. L. and Shleifer, A. (2002). Legal origins. *The Quarterly Journal of Economics*, 117(4):1193–1229.

Greif, A. and Kingston, C. (2011). Institutions: Rules or equilibria? In *Political economy of institutions, democracy and voting*, pages 13–43. Springer.

Greif, A. and Laitin, D. D. (2004). A theory of endogenous institutional change. *American Political Science Review*, 98(4):633–652.

Greif, A. and Tabellini, G. (2010). Cultural and institutional bifurcation: China and Europe compared. *American Economic Review*, 100(2):135–140.

Guo, Y. (2016). Dynamic delegation of experimentation. *American Economic Review*, 106(8):1969–2008.

Guo, Y. and Roesler, A.-K. (2016). Private learning and exit decisions in collaboration. Northwestern University Working Paper.

Halac, M., Kartik, N., and Liu, Q. (2016a). Contests for experimentation. *Journal of Political Economy*.

Halac, M., Kartik, N., and Liu, Q. (2016b). Optimal contracts for experimentation. *The Review of Economic Studies*, 83(3):1040–1091.

Heidhues, P., Rady, S., and Strack, P. (2015). Strategic experimentation with private payoffs. *Journal of Economic Theory*, 159:531 – 551.

Hörner, J., Klein, N. A., and Rady, S. (2015). Strongly symmetric equilibria in bandit games.

Hörner, J. and Samuelson, L. (2013). Incentives for experimenting agents. *The RAND Journal of Economics*, 44(4):632–663.

Hörner, J. and Skrzypacz, A. (2016). Learning, experimentation and information design.

Jha, S. (2013). Trade, institutions, and ethnic tolerance: Evidence from South Asia. *The American Political Science Review*, 107(4):806–832.

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

Keller, G. and Rady, S. (2010). Strategic experimentation with Poisson bandits. *Theoretical Economics*, 5(2):275–311.

Keller, G. and Rady, S. (2015). Breakdowns. *Theoretical Economics*, 10(1):175–202.

Keller, G., Rady, S., and Cripps, M. (2005). Strategic experimentation with exponential bandits. *Econometrica*, 73(1):39–68.

Kingston, C. and Caballero, G. (2009). Comparing theories of institutional change. *Journal of Institutional Economics*, 5(2):151–180.

Klein, N. and Rady, S. (2011). Negatively correlated bandits. *The Review of Economic Studies*, 78(2):693–732.

Kremer, I., Mansour, Y., and Perry, M. (2014). Implementing the "wisdom of the crowd". *Journal of Political Economy*, 122(5):988–1012.

Krusell, P. and Rios-Rull, J.-V. (1996). Vested interests in a positive theory of stagnation and growth. *The Review of Economic Studies*, 63(215).

Lee, I. H. (1993). On the convergence of informational cascades. *Journal of Economic Theory*, 61(2):395–411.

Mansour, Y., Slivkins, A., and Syrgkanis, V. (2015). Bayesian incentive-compatible bandit exploration. In *Proceedings of the Sixteenth ACM Conference on Economics and Computation*, EC '15, pages 565–582, New York, NY, USA. ACM.

Matějka, F. and McKay, A. (2015). Rational inattention to discrete choices: A new foundation for the multinomial logit model. *American Economic Review*, 105(1):272–98.

Mathevet, L., Perego, J., and Taneva, I. (2017). On information design in games. Working Paper.

Matysková, L. (2018). Bayesian persuasion with costly information acquisition. Working paper, CERGE-EI.

McAfee, R. P. and Reny, P. J. (1992). Correlated information and mechanism design. *Econometrica*, 60(2):395–421.

Michalopoulos, S. and Papaioannou, E. (2013). Pre-colonial ethnic institutions and contemporary African development. *Econometrica*, 81(1):113–152.

Milgrom, P. R. (1981). Rational expectations, information acquisition, and competitive bidding. *Econometrica*, 49(4):921–943.

Montada, L. and Lerner, M. J. (1998). *Responses to victimizations and belief in a just world*. Critical issues in social justice. Plenum Press, New York.

Moscarini, G., Ottaviani, M., and Smith, L. (1998). Social learning in a changing world. *Economic Theory*, 11(3):657–665.

Moscarini, G. and Smith, L. (1997). Economic models of social learning. In *Decisions, Games and Markets*, pages 265–298. Springer.

Murto, P. and Välimäki, J. (2011). Learning and information aggregation in an exit game. *The Review of Economic Studies*, 78(4):1426–1461.

Murto, P. and Välimäki, J. (2013). Delay and information aggregation in stopping games with private information. *Journal of Economic Theory*, 148(6):2404 – 2435.

Myerson, R. B. (1986). Multistage games with communication. *Econometrica*, 54(2):323–358.

North, D. C. (2006). *Understanding the process of economic change*. Academic Foundation.

Nunn, N. (2007). Historical legacies: A model linking Africa's past to its current underdevelopment. *Journal of Development Economics*, 83(1):157–175.

Nunn, N. (2008). The long-term effects of Africa's slave trades. *The Quarterly Journal of Economics*, 123(1):139–176.

Nunn, N. (2009). The importance of history for economic development.

Nunn, N. and Wantchekon, L. (2011). The slave trade and the origins of mistrust in Africa. *American Economic Review*, 101(7):3221–3252.

Padró i Miquel, G. (2007). The control of politicians in divided societies: The politics of fear. *The Review of Economic Studies*, 74(4):1259–1274.

Papanastasiou, Y., Bimpikis, K., and Savva, N. (2017). Crowdsourcing exploration. *Management Science*.

Piketty, T. (1995). Social mobility and redistributive politics. *The Quarterly Journal of Economics*, 110(3).

Presman, E. L. (1991). Poisson version of the two-armed bandit problem with discounting. *Theory of Probability and its Applications*, 35(2):307–11.

Rahman, D. (2014). The power of communication. *The American Economic Review*, 104(11):3737–3751.

Rieke, A., Yu, H., Robinson, D., and von Hoboken, J. (2016). Data brokers in an open society. Technical report, Open Society Foundations.

Roesler, A.-K. and Szentes, B. (2017). Buyer-optimal learning and monopoly pricing. *American Economic Review*, 107(7):2072–80.

Roland, G. (2004). Understanding institutional change: Fast-moving and slow-moving institutions. *Studies in Comparative International Development*, 38(4):109–131.

Rosenberg, D., Salomon, A., and Vieille, N. (2013). On games of strategic experimentation. *Games and Economic Behavior*, 82:31 – 51.

Rosenberg, D., Solan, E., and Vieille, N. (2007). Social learning in one-arm bandit problems. *Econometrica*, 75(6):1591–1611.

Rothschild, M. (1974). A two-armed bandit theory of market pricing. *Journal of Economic Theory*, 9(2):185 – 202.

Smith, L. and Sørensen, P. (2000). Pathological outcomes of observational learning. *Econometrica*, 68(2):371–398.

Smith, L., Sørensen, P., and Tian, J. (2017). Informational herding, optimal experimentation, and contrarianism. Technical report.

Smith, L. and Sørensen, P. N. (2011). Observational learning. In Durlauf, S. N. and Blume, L. E., editors, *The New Palgrave Dictionary of Economics*. Palgrave Macmillan, Basingstoke.

Smolin, A. (2015). Optimal feedback design. Working Paper.

Strulovici, B. (2010). Learning while voting: determinants of collective experimentation. *Econometrica*, 78(3):933–971.

Strulovici, B. and Siegel, R. (2018). Judicial mechanism design. Working paper.

Sugaya, T. and Wolitzky, A. (2017). Bounding equilibrium payoffs in repeated games with private monitoring. *Theoretical Economics*, 12(2):691–729.

Sugaya, T. and Yamamoto, Y. (2015). The folk theorem in repeated games with individual learning. Working Paper.

Taneva, I. A. (2016). Information design. Working Paper.

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

Vives, X. (1993). How fast do rational agents learn? *The Review of Economic Studies*, 60(203).

Voigtländer, N. and Voth, H.-J. (2012). Persecution perpetuated: The medieval origins of anti-Semitic violence in Nazi Germany. *The Quarterly Journal of Economics*, 127(3):1339–1392.

Wiseman, T. (2005). A partial folk theorem for games with unknown payoff distributions. *Econometrica*, 73(2):629–645.

Wiseman, T. (2012). A partial folk theorem for games with private learning. *Theoretical Economics*, 7(2):217–239.

Yamamoto, Y. (2014). Individual learning and cooperation in noisy repeated games. *Review of Economic Studies*, 81(1):473 – 500.