Essays in dynamic games of information and risk sharing

ESSAYS IN DYNAMIC GAMES OF INFORMATION AND RISK SHARING

Angelos Diamantopoulos

A thesis submitted for the degree of Doctor of Philosophy

Department of Economics, Mathematics and Statistics, Birkbeck College
University of London

August 2022

Acknowledgements

I would like to thank my supervisor Dr. Emanuela Sciubba for her help and advice. Also, of paramount importance towards the completion of this thesis has been my second supervisor and co-author Dr. Arina Nikandrova, whose help and support has been immense. Finally, there have been many people throughout the years that have helped and inspired me, but for brevity I would like to mention my friend, now Dr., Alkis Georgiadis-Harris and my partner Dr. Sara Stevano, for playing different but significant roles in this process.

Abstract

This thesis consists of three chapters on dynamic games. In the first chapter we consider a dynamic model where a principal delegates learning about the unknown binary state of the world to a biased expert. We find that when preferences of the principal and the expert are sufficiently closely aligned, retaining some decision making authority may be detrimental for the principal. The second chapter develops a model that captures risk taking behaviour of banks. We characterise the unique NE in which banks endogenise the systemic consequence of their actions. In the third chapter we create an algorithm to simulate a risk sharing agreement between two parties that both produce a perishable good. Production is privately known and the parties cannot commit to a long-term contract. We estimate the expected gains of this agreement compared to the case of no sharing. The results can be used to quantify the opportunity cost of not setting a bilaterally-trusted independent authority that monitors and publicly reports electricity generation between two developing countries that want to share electricity.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

This thesis consists of three largely unrelated chapters on dynamic games. The overarching themes of those being information acquisition and risk sharing. All games model environments of incomplete information. Chapter 2 is in continuous time, while the rest are in discrete time. Another connecting thread of this work is that chapters 3 and 4 involve agents who share risk, while chapter 2 involves information sharing in a principal-agent setting.

The second chapter considers a dynamic model where a principal delegates learning about the unknown binary state of the world to a biased expert. Conditional on the state of the world, both the principal and the expert agree on the optimal decision, but at either state of the world, the expert obtains a higher payoff than the principal from one of the decisions. Before taking the decision, the expert may undertake costly learning from two alternative information sources. The principal observes the expert's learning effort and outcomes, but cannot make (contingent) transfers to incentivise learning. Markov equilibria in two situations are compared: when the principal retains the right to terminate the expert's learning to take a decision and when she surrenders completely her decision making authority. When preferences of the principal and the expert are sufficiently closely aligned, retaining some decision making authority may be detrimental for the principal.

The third chapter develops a model that captures risk taking behaviour of banks in a small economy. Systemic risk is defined as the risk of the banking system collapsing due to individual banks' risk taking decisions. The model assumes that there is an institution, such as a Central Bank (CB), with the capacity to bail out any and all banks. I characterize the conditions under which banks internalize systemic risk through investing in mutual assets in order to force the CB into providing liquidity to any bank should it be in distress. In the model, I characterise a unique Nash Equilibrium, in which risk taking by individual banks, through common sharing of investments, occurs in order to maximize system risk. This result depends on the likelihood of distress, the spread of returns across banks and a relatively strict regulator rule.

Chapter 4 extends Hertel (2004)'s model of risk-sharing under one-sided uncertainty to a risk-sharing model under two-sided uncertainty. The main contribution of the chapter is to develop a numerical algorithm to simulate the optimal risk sharing in the absence of commitment to long term contracts. The purpose is to establish whether relational risk sharing is feasible and characterise the necessary parameters for feasibility. The simulations of optimal risk sharing between two parties are compared with the autarky extremes to identify which parameters impact the sustainability of the relational agreements most. The numerical algorithm can be used to assess the extent to which parties that are interested in pooling their resources to share risk would benefit from engaging a trusted third party able to alleviate the information asymmetry.

Chapter 2 is joint work with Dr Arina Nikandrova; the remaining chapters are single authored. My contribution to the co-authored chapter is in the formal proofs of all the results as well as drafting of the paper.

Chapter 2

Delegation of Learning from Multiple Sources of Information *with A. Nikandrova*

2.1 Introduction

Many decisions depend on figuring out the state of the world, but it can often be the case that learning about the state is inaccessible to the decision maker. Instead, learning may require engaging an expert who has access to some learning technology. The expert may be better equipped for undertaking learning due to specialist knowledge or know-how that he has accumulated over many years or because the expert has the exclusive access to key information sources. If the expert and the decision maker's incentives are perfectly aligned, then the latter can hire the former to learn about the state and also to make a decision if and when the expert finds fit. However, if the interests of the decision maker and the expert are not perfectly aligned, it is natural to assume that the decision maker would benefit from retaining at least some decision making authority, particularly if the expert cannot misreport learning outcomes. In this paper, we show that this intuition is not entirely correct. When the decision maker cannot provide monetary rewards for learning, limiting the expert's decision making authority may influence how the expert learns to the detriment of the decision maker.

There are many situations in which the expert is somewhat biased towards an action. For example, consider a financial consultant advising on whether to sell or buy a particular stock. For the most part, it is reasonable to expect that the consultant wants to make the right decision given the current sentiment of the stock market. Nevertheless, he may have a bias towards one of the two actions, stemming perhaps from the risk exposure of his overall portfolio. Alternatively, think of a real estate agent investigating the values of two properties on behalf of a potential buyer. The estate agent may only be interested in the sale commission and hence prefer that the buyer chooses the more expensive property instead of the property which possesses some qualities which the buyer privately values but which are not reflected in the price. As another example consider a government which is interested in learning about the severity of a new disease in order to decide whether to invest in a vaccine technology or whether to invest in controlling the spread of the disease through non-pharmaceutical means. The government may hire an expert, or a group of experts, to investigate the disease. The experts, however, may have a small bias towards the vaccine being developed as this is a field they are more knowledgeable about. In all these interactions, the principal has to engage an expert who is known to be biased. Hence, an important question arises: Would the principal, in addition to delegating learning, ever want to surrender completely the decision making authority?

This paper studies a dynamic model of learning by an expert (he) on behalf of a principal (she). The players are interested in learning about the prevailing binary state of the world as it determines the optimal binary decision. Conditional on the state of the world, both the principal and the expert agree on the optimal decision, but in either state, the expert obtains a higher payoff than the principal from one of the decisions; that is, the expert is biased towards a particular decision.

The principal delegates learning and decision making to the expert. The principal cannot undertake learning herself as only the expert has access to two alternative sources of information. Each source of information generates conclusive good news about one of the states of the world according to a Poisson process. The expert must decide how to divide attention between the two sources. The principal observes the learning outcomes and also observes but cannot directly affect how the expert splits his attention between the information sources.

The principal can choose between two delegation options: complete delegation and flexible delegation. Under **complete delegation**, she commits to surrender completely the authority to make decisions; under **flexible delegation**, she retains the right to take an action as and when she pleases while the expert is still learning.

Our main result states that if preferences of the principal and the expert are sufficiently closely aligned, the principal may find it optimal to completely surrender the authority to make the final decision, instead of retaining the right to intervene. This, seemingly counter-intuitive, result is a consequence of the expert's strategic response when he expects the principal to intervene and terminate learning prematurely by making a decision.

Under flexible delegation the expert learns and acts according to his preferences. However, since preferences of the principal and the expert are misaligned, at some beliefs they disagree on the optimal learning policy.[1] *Inter alia*, the misalignment of preferences means that sometimes the expert finds it optimal to continue learning, when the principal prefers to take the decision immediately and sometimes he may prefer learning from an information source that is sub-optimal from the principal's perspective. Flexible delegation can prevent the expert from prolonged learning, but cannot ensure the right type of learning.

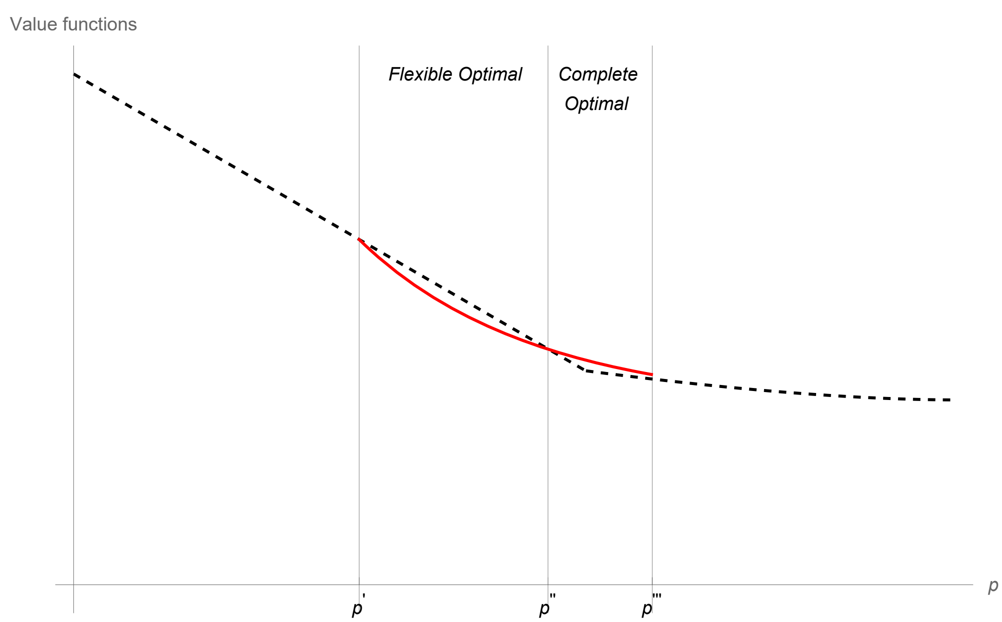


Figure 2.1: **Main result.** The red curve corresponds to the principal's value function under complete delegation; the black curve corresponds to the principal's value function under flexible delegation. The principal prefers complete delegation whenever the red curve lies above the black curve.

[1]The nature of their disagreement is discussed further in Section 2.6.

Figure 2.1 demonstrates that there is a non-empty belief interval in which the expert prefers learning, while the principal would derive a higher utility by taking an action immediately. In this region (from p' until p''), the principal's utility given the experts optimal learning (red line) is lower than the principal's utility from an immediate action (black line). It follows immediately that if the prior belief about the state falls within this region, the principal would always favour flexible to complete delegation.

It is less immediate that the principal does not always want to retain the right to intervene. In particular, as seen in Figure 2.1 there is a belief interval (from p'' until p''') in which the principal's utility when the expert learns optimally under *flexible delegation* (red line) is higher than the utility the principal derives from *complete delegation* (black line). Hence, in this region, the principal is better off relinquishing her decision making rights. This region arises because flexible delegation cannot prevent the expert from learning from the sub-optimal source. In fact, flexible delegation may make expert's choice of information sources even more sub-optimal for the principal than it is under complete delegation. Knowing that under flexible delegation, the principal will want to terminate expert's learning prematurely, the expert may choose information sources in a way that delays or even avoids termination of learning by the principal. As delay is costly to the principal, under certain conditions, the principal is better off allowing the expert to learn freely.

The main contribution of our paper is to establish that paradoxically, flexible delegation may be sub-optimal for the principal and to characterize the conditions under which complete delegation is optimal. To this end, we guess and verify the equilibrium strategies of both players under complete and under flexible delegation and then show that, given these equilibrium strategies, there exist regions in which complete delegation is optimal for the principal.

The rest of the paper is organized as follows. In Section 2.2 we discuss the related literature and situate our paper therein. Section 2.3 describes the setup. Section 2.4 outlines the main result of the paper and provides a road map for its proof. Section 2.6 describes the principal's first-best learning strategy. Section 2.6 characterizes the equilibrium under complete delegation. Section 2.7 fully characterizes the equilibrium under flexible delegation. Section 2.8 extends the discussion of our main result. Section 2.9 concludes the paper.

2.2 Literature Review

Our paper is directly related to the literature on optimal dynamic learning and to various strands of the delegation literature.

Optimal dynamic learning

The literature on optimal dynamic learning, pioneered by Wald (1947), has attracted a lot of attention recently from economic theorists. Optimal information acquisition in continuous time has been studied comprehensively by Zhong (2022). In their paper, the decision maker (DM) is allowed to choose any dynamic signal process in order to learn about the state and act accordingly, subject to a informativeness cost. They show that the DM will choose a Poisson signal process. We assume from the outset that the expert has access to two exogenous Poisson information processes and optimally chooses between them.

An alternative approach to modelling optimal learning is taken in Ke and Villas-Boas (2017), in which a DM is deciding among many alternatives and can purchase informative signals for each of those. In our paper, the principal has to choose between two alternatives actions and the validity of each action can be confirmed by an information source.

Our paper builds on the sequential decision problem of Che and Mierendorff (2019) with two sources generating information about two mutually exclusive states of the world. Our contribution is to embed the decision problem from Che and Mierendorff (2019) into a game with two players, splitting control of learning and final decision between two players, the principal and the expert.

Mayskaya (2020) characterizes the unique optimal learning strategy in a decision problem with two information sources, as in Che and Mierendorff (2019), but with one additional unfindable state.[2] She finds that the optimal learning strategy has two phases. In the first phase, the decision maker optimally learns from the source that could confirm a more plausible state of the world; in the second phase, the decision

[2]Nikandrova and Pancs (2018) also study a decision problem, but with four states.

maker focuses on learning about the least likely (findable) state of the world. The decision problem of Che and Mierendorff (2019) also features such phases and a lot of our analysis focuses on examining the consequences of the misalignment of principal and expert's learning phases.

Our paper is also related to Brocas and Carrillo (2007) who analyse a game in which a leader collects information about a binary world and a follower can then use this information to take actions that impact both agents. The paper describes the leader's stopping rule and his equilibrium information rents. In our paper, the principal may decide when to terminate the expert's learning, but we ask a normative question of whether such interference is optimal for the principal.

Delegation of decision making

The literature on delegation of decision making authority mostly focuses on agents who are more informed than the principal. In such cases, the principal faces a challenge of inducing the biased agent to act in the principal's interest. Thus, Alonso and Matouschek (2008) consider a model in which the principal offers a menu of decisions that an informed agent can make. The authors derive the conditions under which an interval allocation of decisions is optimal for the principal. In our model, the principal faces a challenge of inducing the biased expert to learn in a way the principal would want to learn.

In another strand of the literature, the principal would like to elicit truthful information from the biased agent. Thus, Dessein (2002) find that due to the cost of communication, the principal finds it optimal to delegate decision making authority to the agent when the agent's bias is small enough. Although for different reasons, this resembles our main finding that the principal wants to relinquish her ability to intervene. In contrast, in a cheap talk setting, Argenziano, Severinov, and Squintani (2016) compare implications of delegating versus communicating of learning between a biased expert and a principal. They find that for low bias levels, communication is preferable to delegation of learning and decision making to the expert, both when the expert's learning is observed and unobserved. The intuition for the latter is that as the expert cannot credibly communicate the amount of learning deviations from equilibrium path can be costly for the expert.[3] Our paper contributes to the literature on delegation by considering a different dimension. Rather than trying to elicit agent's private information, we consider a set up with common knowledge but where the agent alone has access to a learning technology. We ask the question whether in the absence of money transfers and long-term contracting, the principal would want to fully delegate the decision making authority to a biased expert.

Delegation of learning

A central theme within the literature of delegation of learning is that of observability. Much of the literature focuses on unobservable learning by a biased agent who discloses their findings to the principal. In this literature the issues of truth revelation and optimal learning are key. For example, Escobar and Zhang (2021), study a dynamic delegation model in which an agent learns privately about the profitability of an investment opportunity. They characterise the optimal truth-revealing contract, and find that it may require the principal to delay action, in order for the agent to be allowed to learn freely. On the other hand, Herresthal (2022) is interested in the trade-off between observable and unobservable learning. They consider a framework in which a DM delegates learning via sequential experimentation about the safety of a product to an agent with lower safety standards than her, meaning that the agent is willing to retain the product at a lower certainty threshold. They show that both the agent and the DM can be strictly better off if testing is privately known by the agent as opposed to publicly known if their difference in safety standards is above some threshold. The intuition behind this result is that when learning is observable the agent can learn strategically, i.e. stop or continue experimenting based on the DM's posterior. Conversely, when learning is private and the agent can withhold outcomes, the burden of proof is stronger and so more evidence gathering is required to convince the DM. In our model, the principal observes, but cannot directly control the allocation of the expert's learning effort. Therefore we are not focusing on the revelation of learning outcomes, rather we are focused on learning that is observable and in situations in which both the principal and the agent can take decisions.

Garfagnini (2011) consider an experimentation game between a principal and biased agent. Differently from our set up, they analyse an exponential bandit model with

[3] Deimen and Szalay (2015) also compare delegation of decision-rights and communication in a sender-receiver game with endogenous information.

a risky and a safe arm, similar to Keller, Rady, and Cripps (2005). In their setting the principal decides which actions is taken at every interval while the agent controls the rate of learning. The agent is biased toward the risky action being taken, and so she has an incentive to delay information so that the bad news do not arrive. Interestingly this effect dissipates when effort is unobservable.

Guo (2016) characterises the optimal contract when a principal delegates experimentation, modelled as a two-armed bandit with a Poisson arm of unknown intensity, to a biased, privately informed agent. The optimal contract allows the expert to experiment fully as long as the principal's belief remains above a threshold.

2.3 Model

We consider a game between two players, a principal (she) and an expert (he). Time is continuous and indexed by $t \geq 0$. The time horizon is infinite. Both players discount future exponentially at rate $r > 0$.

Problem The players are interested in taking the decision that is optimal for the unknown state of the world. There are two states of the world, $\omega \in \{A, B\}$ and two decisions $x \in \{a, b\}$.

Delegation The expert, but not the principal has, access to an information technology that enables learning about ω. At any time $t > 0$, the principal delegates learning as well as decision making to the expert thereby giving the expert authority to learn about ω and to take final decision $x \in \{a, b\}$. The principal observes but cannot directly affect (through contingent payments or coercion) learning and the final decision taken by the expert.

We distinguish two types of delegation: *complete delegation* and *flexible delegation*. Under complete delegation, the principal surrenders completely the authority to make decisions to the expert. Under flexible delegation, at each time, t, while the expert has not taken action x yet, the principal retains the right to override the expert and take decision $y \in \{a, b\}$ herself.

Learning At each time t, when he is given the authority to learn, the expert allocates a unit of learning intensity between two information sources, A and B. Thus, if the expert devotes a fraction $\alpha \in [0,1]$ of his learning intensity to source A, then the learning intensity devoted to source B is $1-\alpha$. Information source A reveals conclusively state A; that is, when the expert learns from A with intensity α, source A generates a signal with a Poisson arrival rate α if $\omega = A$ and no signal if $\omega = B$. Symmetrically, information source B reveals conclusively state B; that is, when the expert learns from B with intensity $1-\alpha$, source B generates a Poisson signal with arrival rate $(1-\alpha)$ if $\omega = B$ and no signal otherwise. The learning process of the expert is denoted by $(\alpha_t)_{t\in\mathbb{R}_+}$.

At $t = 0$, the principal and the expert share a common prior belief p_0 that the state is A. Subsequently, the principal observes $(\alpha_t)_{t\in\mathbb{R}_+}$ as well as learning outcomes. Hence, the belief that the state is A remains common throughout the game. Let p_t denote the belief at time t. By Bayes' rule, as long as the expert's learning generates no signal, p_t evolves according to[4]

$$\dot{p}_t = -\left(2\alpha_t - 1\right) p_t \left(1 - p_t\right). \tag{2.1}$$

Payoffs Let u_x^ω and v_x^ω denote the utility of the principal and the expert respectively, conditional on decision x and state ω. The states are labeled so that both the principal and the expert strictly prefer taking the decision that matches the state, that is, $u_a^A > \max\left\{0, u_b^A\right\}$, $u_b^B > \max\left\{0, u_a^B\right\}$, $v_a^A > \max\left\{0, v_b^A\right\}$ and $v_b^B > \max\left\{0, v_a^B\right\}$. Moreover, independently of the state, the expert prefers decision a relative to the principal and the principal prefers decision b relative to the expert. In particular, for $\Delta \geq 0$, the payoff profiles (u_x^ω, v_x^ω) are

P, E	A	B
a	$\rho,\ \rho+\Delta$	$-\rho,\ \Delta-\rho$
b	$\Delta-\rho,\ -\rho$	$\rho+\Delta,\ \rho$

where $\rho > \Delta \geq 0$. It is without loss to take $\rho = 1$ and so the paper proceeds thus.

Throughout the paper, we assume that the principal's and expert's preferences are not too misaligned.

Restriction 2.1. $\Delta \leq \max\left\{\bar{\Delta}(r), 1\right\}$, *where* $\bar{\Delta}(r) = \frac{\sqrt{36r^2+28r+1}-6r-1}{4r} > 0$

This restriction ensures that the principal finds it worthwhile to delegate learning to

[4] See Appendix 2.A.1 for derivation.

the expert. [5]

Given belief p, let

$$U_x(p) \equiv pu_x^A + (1-p)\,u_x^B \text{ and } V_x(p) \equiv pv_x^A + (1-p)\,v_x^B \tag{2.2}$$

denote the expected payoff of the principal and the expert, respectively, from taking decision x immediately.

Timing Over an infinitesimal time interval of length $h > 0$, a heuristic timeline of the stage game between the principal and the expert is:

1. The expert decides whether to take the final decision $x \in \{a, b\}$;
2. If the expert does not take the final decision, he chooses learning intensity α_t;
3. Having observed α_t, the principal decides whether to take the final decision $y \in \{a, b\}$;
4. If the principal has not taken the final decision, the learning outcomes are publicly observed and belief p_t is updated according to Bayes rule.

Equilibrium Concept We focus our attention to Markov perfect equilibria. A Markov strategy for the expert is a tuple (α, X), where learning strategy $\alpha : [0,1] \to [0,1]$ maps a belief p into learning intensity $\alpha(p) \in [0,1]$, while decision strategy $X : [0,1] \to \{a, b, \emptyset\}$ maps p onto the final decision $X(p) \in \{a, b, \emptyset\}$.[6] A Markov strategy for the principal $Y : [0,1]^2 \to \{a, b, \emptyset\}$ maps a belief p and expert's learning intensity α onto the decision $Y \in \{a, b, \emptyset\}$.

A strategy profile $\sigma = \Big(\big(\alpha, X\big), Y\Big)$ induces a Markov learning policy $\Big(\alpha_t, X_t, Y_t\Big) = \left(\alpha\big(p_t\big), X\big(p_t\big), Y\Big(p_t, \alpha\big(p_t\big)\Big)\right)$, the belief process $\{p_t^\sigma \mid t \geq 0\}$ and the stopping time τ that designates when all learning stops and a final decision is taken. The stopping time τ is the minimum of stopping times τ^P and τ^E, where τ^P designates when the principal takes the final decision and τ^E designates when the expert takes the final decision.

[5] The genesis of this restriction is discussed in more detail in Section 2.6. The restriction is somewhat stronger than assuming that when the expert optimally stops learning, he and the principal agree on the optimal action. This stronger restriction is helpful for the full characterization of the equilibrium as can be seen in 2.A.15.

[6] $X(p) = \emptyset$ indicates that the expert does not take a final decision at p.

A strategy profile σ and an initial belief p_0 induce expert's expected discounted payoff:

$$I(p,\sigma) \equiv \mathbb{E}\left[\mathbb{I}\left\{\tau = \tau^E\right\} e^{-r\tau} \max\left\{V_a\left(p_\tau^\sigma\right), V_b\left(p_\tau^\sigma\right)\right\} \right.$$
$$\left. +\mathbb{I}\left\{\tau = \tau^P\right\} e^{-r\tau} V_{Y_\tau}\left(p_\tau^\sigma\right) \mid p^\sigma(0) = p\right], \tag{2.3}$$

where $\mathbb{I}\{\cdot\}$ is an indicator function and the expectation is with respect to the induced belief process $\{p_t^\sigma \mid t \geq 0\}$. Similarly, a strategy profile σ and an initial belief p induce principal's expected discounted payoff:

$$J(p,\sigma) \equiv \mathbb{E}\left[\mathbb{I}\left\{\tau = \tau^E\right\} e^{-r\tau} U_{X_\tau}\left(p_\tau^\sigma\right), \right.$$
$$\left. +\mathbb{I}\left\{\tau = \tau^P\right\} e^{-r\tau} \max\left\{U_a\left(p_\tau^\sigma\right), U_b\left(p_\tau^\sigma\right)\right\} \mid p_0^\sigma = p\right]. \tag{2.4}$$

An equilibrium strategy profile $\sigma^* = ((\alpha^*, X^*), Y^*)$ for all p satisfies

$$V(p) \equiv \sup_{(\alpha,X)} I(p, ((\alpha, X), Y^*)) \tag{2.5}$$

and

$$U(p) \equiv \sup_{Y} J(p, ((\alpha^*, X^*), Y)), \tag{2.6}$$

where the maximization is over all admissible Markov strategies, $U(p)$ is the principal's value function and $V(p)$ is the expert's value function.

2.4 Main Result

Theorem 2.1 states the main result of the paper.

Theorem 2.1. *Fix $r \leq \bar{r}(0)$. There exists a threshold $\underline{\Delta}(r)$ such that*

- *for $\Delta \leq \underline{\Delta}(r)$, the principal is better off under complete delegation than under flexible delegation for some p;*
- *for $\Delta > \underline{\Delta}(r)$, principal is better off under flexible delegation for any p.*

Proof. See Appendix 2.A.24. $\square$

To build up to proving the main result, we first briefly discuss the first best learning strategy that the principal would pursue if she had access to expert's learning technology. We then contrast principal's first best learning strategy to expert's first-best learning strategy. Expert's first best learning strategy plays a dual role in the analysis. First, it constitutes equilibrium learning strategy under complete delegation when the expert is learning and acting optimally in order to maximise his utility. Second, expert's learning strategy is a stepping stone towards deriving the equilibrium strategies under flexible delegation, where the principal retains the right to intervene and take an action while the expert is still learning. Once the equilibrium learning strategies under complete and under flexible delegation are characterised, we prove Theorem 2.1 by directly comparing the principal's payoff under complete and flexible delegation.

Throughout the paper, quantities without any superscript correspond to flexible delegation, superscript E indicates quantities under complete delegation and superscript P indicates principal's first best learning when the principal has access to the same learning technology as the expert.

2.5 First-Best Benchmark

From the perspective of the principal, the first-best outcome obtains when the principal has access to the same learning technology as the expert. In this case, the principal must solve a decision problem. The solution to this decision problem is characterised in Che and Mierendorff (2019). To keep the analysis self-contained, in this section, we describe (but do not prove) the solution to the principal's problem.

The principal, having access to the learning technology, chooses a learning strategy (α_t) as well as a stopping time τ, at which point she takes action a or b in order to maximize her discounted expected utility. Abusing the notation somewhat, her action space remains $Y = \{a, b, \emptyset\}$. She is therefore facing the following maximization problem

$$U^P(p) \equiv \max_{\alpha,\tau} \mathbb{E}\left[e^{-r\tau} \max\left\{U_a\left(p_\tau^\alpha\right), U_b\left(p_\tau^\alpha\right)\right\} \mid p_0^\alpha = p\right], \tag{2.7}$$

subject to belief updating equation (2.1) and given the boundary condition

$$U^P(p) = U_a(p) \vee U_b(p). \tag{2.8}$$

$$\Big|_{p=0}\ \underbrace{\overline{\qquad\qquad\qquad}}_{\text{decision } b}\ \underline{p}^{*P}\ \underbrace{\leftarrow\leftarrow\leftarrow\leftarrow}_{\alpha=1}\ \check{p}^{P}\ \underbrace{\rightarrow\rightarrow\rightarrow\rightarrow}_{\alpha=0}\ \bar{p}^{*P}\ \underbrace{\overline{\qquad\qquad\qquad}}_{\text{decision } a}\ \Big|_{p=1}$$

Figure 2.2: Contradictory learning strategy

The Hamilton-Jacobi-Bellman (HJB) equation associated with the problem is

$$rU^P(p) = \max_{\alpha} \left\{ \begin{array}{c} \alpha p \left(u_a^A - U^P(p)\right) + (1-\alpha)(1-p)\left(u_b^B - U^P(p)\right) \\ -(2\alpha - 1)\, p\,(1-p)\, U^{P\prime}(p) \end{array} \right\} \tag{2.9}$$

The HJB equation (2.9) is linear in α and therefore has a bang-bang solution, i.e. it is maximised at either $\alpha = 0$ or $\alpha = 1$, wherever $U^P(p)$ is differentiable.

To characterize the principal's first-best value function consider the payoffs derived by four distinct learning policies. The first is the learning policy splits attention between the sources equally and the principal takes no decision until a signal arrives. One can immediately observe that $\dot{p} = 0$ when $\alpha = 1/2$, hence, conditionally on no signal arriving, posterior beliefs remain unchanged at $p_t = p_0, \forall t$. Hence, under this policy the principal's payoff would equal the discounted expected payoff of taking the decision that matches the state. We denote principal's payoff from this policy by $U^*(p)$. [7]

Next, let $U_A^P(p)$ denote the principal's payoff when she is focusing all her attention on source A, and makes decision $Y = b$ when the utility of immediate decision $Y = b$ exceeds that of learning from source A, conditional on no signal arriving. While the principal learns from source A and sets $\alpha = 1$, according to belief updating equation (2.1), in the absence of the signal, she becomes progressively more pessimistic that the state is A, which makes decision $Y = b$ more attractive. Similarly, let $U_B^P(p)$ denote payoff from learning from source B only, i.e. $\alpha = 0$, and taking decision a optimally in the absence of a signal. When principal learns from source B and signal does not arrive, she becomes progressively more pessimistic that the state is B, which makes decision $Y = a$ more attractive.[8]

We call *contradictory learning* the upper envelope of the payoff derived from learning from either only source A or only source B until it is optimal to stop. That is,

[7] For analytical derivation see Appendix 2.A.2

[8] For analytical derivations of these functions see Appendix 2.A.4

$$\underset{p=0}{|}\;\underbrace{\rightarrow\rightarrow\rightarrow\rightarrow\rightarrow\rightarrow\rightarrow\rightarrow\rightarrow\rightarrow}_{\alpha=0}\;\underset{\alpha=\frac{1}{2}}{p^{*P}}\;\underbrace{\leftarrow\leftarrow\leftarrow\leftarrow\leftarrow\leftarrow\leftarrow\leftarrow\leftarrow\leftarrow}_{\alpha=1}\;\underset{p=1}{|}$$

Figure 2.3: Confirmatory learning strategy

principal's payoff from contradictory learning strategy is

$$U^P_{ct}(p) = \max\left\{U^P_A(p), U^P_B(p)\right\}.$$

The term contradictory learning was coined by Che and Mierendorff (2019). Such learning dictates that full attention is devoted to the source that generates evidence against the currently preferred action. That is, the principal learns from source B if state A is relatively likely, and from source A if state B is relatively likely. In the absence of a signal, the principal's belief drifts in the direction which strengthens the original belief until the immediate decision becomes optimal. Figure 2.2 depicts belief updating under contradictory learning.

Finally, let *confirmatory learning* refer to a learning strategy that aims to confirm the state that the principal currently prefers for immediate action. Formally, there is an optimally chosen threshold belief p^{*P} such that for $p < p^{*P}$, confirmatory learning focuses attention wholly on source B, i.e., $\alpha = 0$. Note that in this case, Bayesian updating dictates that in the absence of a signal, posterior beliefs would become progressively more pessimistic that the state is B. Symmetrically, for beliefs $p > p^{*P}$, the confirmatory learning policy dictates learning from source A, i.e., $\alpha = 1$. Let $U^P_{cf}(p)$ denote principal's payoff from confirmatory learning.[9]

Figure 2.3 demonstrates that under confirmatory learning, starting from any belief $p \neq p^{*P}$, conditional on no signal arriving, the posterior belief would tend towards the absorbing belief p^{*P}. At belief p^{*P}, the confirmatory learning policy dictates that attention is split between source B and A equally, i.e. $\alpha = 1/2$. It follows that at p^{*P}, the payoff from the confirmatory learning strategy is $U^P_{cf}(p^*) = U^{*P}(p^{*P})$, because no learning takes place, beliefs remain unchanged and the principal acts only after the arrival of a signal.

Che and Mierendorff (2019) prove that the upper envelope of the confirmatory and

[9]For analytical derivation see Appendix 2.A.7

$$\underset{p=0}{|}\ \underbrace{\overline{\qquad\qquad\qquad}}_{\text{decision } b}\ \underline{p}^{*P}\ \overbrace{\underbrace{\leftarrow\leftarrow\leftarrow\leftarrow\leftarrow}_{\alpha=1}}^{\text{contradictory}}\ \underline{p}^{P}\ \overbrace{\underbrace{\rightarrow\rightarrow\rightarrow}_{\alpha=0}\ \underset{\alpha=\frac{1}{2}}{p^{*P}}\ \underbrace{\leftarrow\leftarrow\leftarrow}_{\alpha=1}}^{\text{confirmatory region}}\ \bar{p}^{P}\ \overbrace{\underbrace{\rightarrow\rightarrow\rightarrow\rightarrow\rightarrow}_{\alpha=0}}^{\text{contradictory}}\ \bar{p}^{*P}\ \underbrace{\overline{\qquad\qquad\qquad}}_{\text{decision } a}\ \underset{p=1}{|}$$

Figure 2.4: Combination of contradictory and confirmatory learning

contradictory learning strategies described above, in conjunction with the boundary condition 2.8, constitute a uniquely optimal strategy for the decision maker. Depending on the discount rate r, which acts as the implicit cost of learning, the optimal strategy is one of three types.

- If r is higher than threshold $\bar{r}^P(\Delta)$, immediate decision is optimal for any p.
- For intermediate values of r, that is, $\bar{r}^P(\Delta) > r \geq \underline{r}^P(\Delta)$, the optimal strategy involves contradictory learning depicted in Figure 2.2.
- For low values of r, that is, $r < \underline{r}^P(\Delta)$, both contradictory and confirmatory learning occur as depicted in Figure 2.4.

Intuitively, the higher r the lower the value of learning, because learning delays the payoff from final decision $x \in \{a, b\}$. When the cost of delay is sufficiently high, that is, r is higher than $\bar{r}^P(\Delta)$, no learning at all is optimal and the principal acts on her prior belief. The threshold $\bar{r}^P(\Delta)$ is the smallest discount rate r such that the principal prefers choosing the action immediately to learning for $\delta \to 0$ units of time and only then choosing the optimal action:[10]

$$r \geq \frac{4-\Delta^2}{2\Delta} \equiv \bar{r}^P(\Delta).$$

Note that

$$\lim_{\Delta \to 0^+} \bar{r}^P(\Delta) = \infty$$

That is, for very small bias Δ the principal prefers some learning.

Confirmatory learning can be optimal only if the cost of delay is sufficiently low. On the one hand, confirmatory evidence arrives at higher expected rate than contradictory evidence. On the other hand, the arrival of confirmatory evidence does not change the decision maker's optimal decision and so acquiring confirmatory evidence for a short

[10]For derivation, see Appendix 2.A.8

amount of time only causes unnecessary delay. However, prolonged period of learning without any confirmatory evidence arriving causes the decision maker to update beliefs sufficiently to potentially sway the optimal decision. Hence, confirmatory learning is optimal only when the decision maker plans for prolonged learning, that is, when the cost of delay is sufficiently low. The highest discount rate under which confirmatory learning is optimal, $\underline{r}^P(\Delta)$, ensures that at p^{*P}, the principal prefers making the right decision after potentially long delay to shorter contradictory learning which nevertheless may lead to a wrong decision.[11] Formally, the threshold value $\underline{r}^P(\Delta)$ is the value of r that makes the below condition hold with equality:

$$\left(\frac{r}{r+2-\Delta(r+1)}\right)^r \geq \frac{1}{2r+1}$$

The inequality above can be re-arranges to obtain the highest level of Δ, as a function of r, for which the optimal strategy includes a confirmatory and contradictory part:

$$\Delta = \frac{-r(2r+1)^{\frac{1}{r}} + r + 2}{r+1}. \tag{2.10}$$

2.6 Complete Delegation

It is straightforward to see that under complete delegation the expert will pursue his first-best learning strategy. Indeed, since the principal cannot intervene, the expert has no reason to pursue a learning strategy different from the one he would pursue if he was facing an individual decision making problem.

Thus, with regards to the expert's value function and optimal strategy, the analysis is similar to the principal's first-best described in the previous section, but adjusted for the expert's payoffs.

In what follows, superscript E denotes the complete delegation case. In addition, to denote various threshold beliefs we use the same notation as before, but substitute superscript P with E. For example, $\underline{p}^{*E}$, denotes the belief threshold below which the expert takes immediate action and above which he pursues a learning strategy with $\alpha = 1$.

[11] For derivation see Appendix 2.A.9

Lemma 2.1 demonstrates that for a given Δ and r, the optimal learning strategy of the principal is of the same type as the optimal learning strategy of the expert. That is, one of the three cases obtain:

Case 1 Both the expert and the principal prefer the immediate decision for any p.

Case 2 Both the expert and the principal would like to follow contradictory learning strategy for p sufficiently far away from the extreme values of 0 and 1.

Case 3 Both the expert and the principal would like to have regions of both contradictory and confirmatory learning.

Lemma 2.1. *For any* Δ, $\bar{r}^E(\Delta) = \bar{r}^P(\Delta) \equiv \bar{r}(\Delta)$ *and* $\underline{r}^E(\Delta) = \underline{r}^P(\Delta) \equiv \underline{r}(\Delta)$.

Proof. The result follows from the symmetry of the payoff matrix. In particular, note that the payoffs for the expert and the principal are diametrically opposite. By this we mean that $u_x^\omega = v_{x^-}^{\omega^-}$, where the negative superscript denotes the alternative state or action. To put it in another way, the payoffs of the principal are identical to those of the expert if one relabels the payoffs of matching state A as those of matching state B and those of mismatching state is A with those of mismatching state B. Lemma 2.1 is therefore equivalent to stating that the labeling of each state does not bear any relevance to the type of the optimal learning strategy, which is self-evident. Simple algebra confirms this to be the case. □

Even though the principal and the expert agree on the type of learning strategy to pursue, due to bias Δ, at a particular belief p, they may disagree on the source of information from which to learn or whether to learn at all. In particular, because the expert prefers decision a relative to the principal irrespective of the state, the expert's learning region is shifted to the left relative to the principal's first-best learning region (Lemma 2.2 demonstrates). Thus, for example, the expert takes immediate decision a or b at lower p than the principal would have liked. Similarly, for the expert all the switches of learning from one source to another are shifted to the left relative to the principal.

Lemma 2.2. *For any* $\Delta > 0$, *the expert's learning region is shifted to the left relative to the principal's learning region. Specifically,* $\underline{p}^{*E} < \underline{p}^{*P}, \bar{p}^{*E} < \bar{p}^{*P}$ *and* $p^{*E} < p^{*P}$.

Proof. See Appendix 2.A.11.

□

Recall that we impose Restriction 2.1, which can equivalently be expressed as a restriction on r for a given Δ, $r < \bar{r}(\Delta)$.[12] This restriction ensures that if the expert find's it optimal to learn forever, then the principal also prefers learning to her outside option, i.e. this restriction ensures that $U^E_{cf}\left(p^{*E}\right) > U_b\left(p^{*E}\right)$. Furthermore, Restriction 2.1 has the direct implication that when the expert stops learning under complete delegation, the principal and the expert agree on the optimal decision and thus the principal derives some value from the expert's learning.[13]

Assuming that Restriction 2.1 holds, Figures 2.5 and 2.6 compare the principal's value under complete delegation to principal's value in the first-best world where the principal has access to the same learning technology as the expert. As a result of disagreement on the optimal learning strategy at a given p, the principal's expected payoff when the expert undertakes learning on his behalf falls short of his first-best expected payoff. Furthermore, the principal's value function under expert's optimal learning is discontinuous at a belief whenever the belief process is certain to move away from that belief. The beliefs with discontinuities are boundaries of the regions where the type of learning switches, but because the belief process moves away from those beliefs, the value matching does not hold.

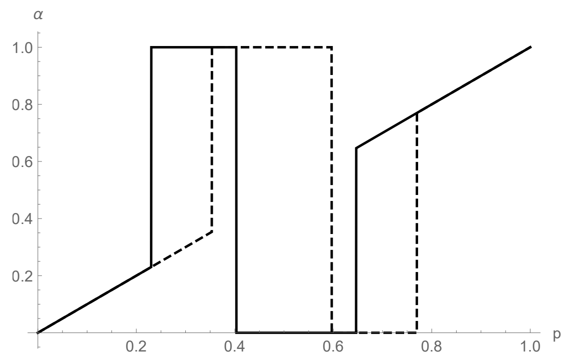
Whenever $r < \bar{r}(\Delta)$, for some not-too-low p, the delegation of learning allows the principal obtain higher payoff than the payoff he obtains from the immediate decision. However, for sufficiently low p the principal prefers an immediate decision to any learning, while the expert finds it optimal to continue learning. Hence, for low p the principal would prefer to take over the final decision making from the expert.

2.7 Flexible Delegation

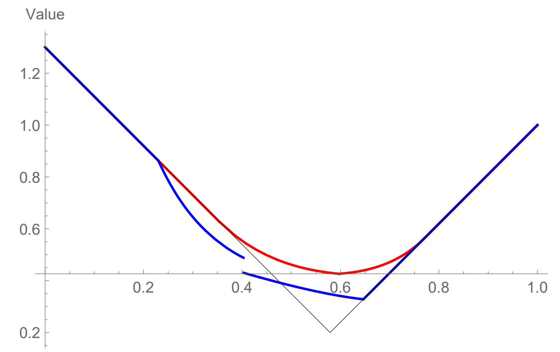
In this section, we derive equilibrium strategies under flexible delegation.

Under flexible delegation, any principal's intervention is necessarily one-sided. For low p, the principal stops contradictory learning from source A at higher p than the expert would have liked. For high p, the principal would also like to learn from source B for higher p than the expert does. The principal, however, cannot do anything about it as she cannot coerce the expert to learn for longer than the expert finds optimal.

[12] For derivation see Appendix 2.A.12.
[13] See Appendix 2.A.13

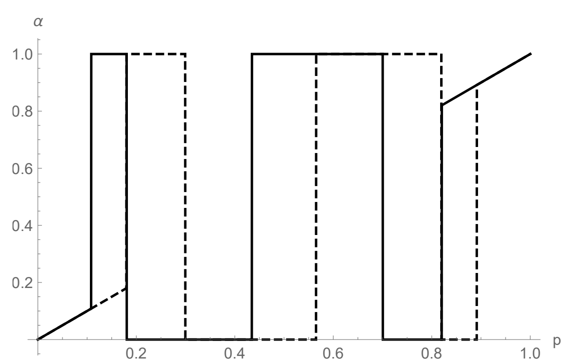


(a) Learning strategy. The optimal learning strategy from the perspective of the principal (dashed line) and the expert (solid line). When the player finds it optimal not to learn, α is set to be equal to belief p.

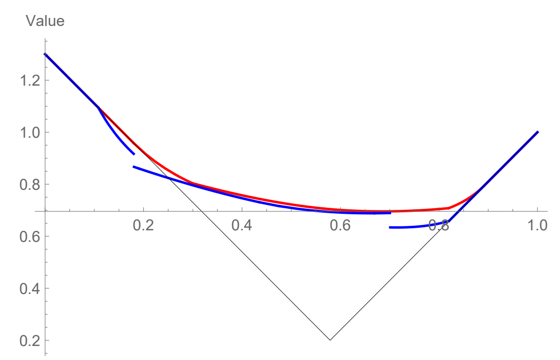


(b) The value function. The red dashed line is the value the principal would obtain if he had access to expert's leaning technology; the blue solid line is the value the principal obtains when he delegates learning and all decision making rights to the expert; the black dotted line is the principal's value from immediate decision.

Figure 2.5: The optimal policy's prescription for each belief in Case 2.



(a) Learning strategy. The optimal learning strategy from the perspective of the principal (dashed line) and the expert (solid line). When the player finds it optimal not to learn, α is set to be equal to belief p.



(b) The value function. The red dashed line is the value the principal would obtain if he had access to expert's leaning technology; the blue solid line is the value the principal obtains when he delegates learning and all decision making rights to the expert; the black dotted line is the principal's value from immediate decision.

Figure 2.6: The optimal policy's prescription for each belief in Case 3.

Armed with this observation, we derive equilibrium learning policy using "guess and verify" approach. Thus, in Section 2.7.4 we guess the equilibrium learning strategy and compute the value functions that this strategy delivers. In Appendix 2.A.23 we show that the HJB equations for both the expert and the principal hold at all points of differentiability of the candidate value functions.[14]

2.7.1 Optimality Conditions

Given strategy of the principal σ^p, the expert solves an optimal control free-boundary problem. By the dynamic programming principle (DPP), the expert's value today equals the expert's expected discounted continuation value at an arbitrary future stopping time plus the expected discounted payoffs accruing until that time. The intervening flow payoffs and the eventual continuation value depend on the intervening learning policy, chosen to maximize expert's value today. The DPP implies that wherever V is differentiable, V can be characterized through the HJB equation:

$$\begin{aligned} rV(p) \geqslant \max_{\alpha} \Big\{ & \mathbb{I}\{Y(p,\alpha)=\emptyset\} \left(p\alpha\left(v_a^A - V(p)\right)\right) + (1-p)(1-\alpha)\left(v_b^B - V(p)\right) \\ & - (2\alpha-1)\,p(1-p)\,V'(p) + \mathbb{I}\{Y(p,\alpha)=a\}\,V_a(p) \\ & + \mathbb{I}\{Y(p,\alpha)=b\}\,V_b(p) \Big\}, \; p \in (0,1) \end{aligned} \tag{2.11}$$

subject to the boundary condition[15]

$$V(p) = V_a(p) \vee V_b(p). \tag{2.12}$$

The expert's continuation payoff depends on the principal's strategy through the principal's ability to terminate learning. The boundary condition ensures that the expert's value function is at least as high as the value that he can obtain from immediate decision $x \in \{a, b\}$.

[14]The value functions, however, have points of non-differentiability and so further technical work has to be undertaken to verify that the derived value functions are indeed solutions to the principal and expert's problems. Since we derive the value functions from the conjectured policies, we remain confident that the points of non-differentiability do not invalidate our solution.

[15]Operator $\vee$ is the binary max operator.

Given strategy of the expert σ^E, the principal faces a stopping problem with value that satisfies the HJB equation

$$\begin{aligned} rU(p) \geq p\alpha(p)\left(u_a^A - U(p)\right) + (1-p)(1-\alpha(p))\left(u_b^B - U(p)\right) \\ - (2\alpha(p) - 1)\, p(1-p)\, U'(p), \end{aligned} \tag{2.13}$$

subject to the boundary condition

$$U(p) = U_a(p) \vee U_b(p). \tag{2.14}$$

In a Markov equilibrium, (2.13) and (2.11) are satisfied simultaneously for all $p \in [0,1]$. Intuitively, in equilibrium, the expert pursues a learning strategy that maximizes his utility subject to preferring learning to immediate action, as well as the principal not terminating learning. Thus, to solve for an equilibrium, we need to find V that satisfies (2.11) subject to (2.12) as well as an additional constraint

$$U(p) \geq U_a(p) \vee U_b(p). \tag{2.15}$$

Constraint (2.15) ensures that on the learning continuation region, the principal's value is at least as high as her value from taking decision $y \in \{a, b\}$ immediately. That is, since the principal can only stop learning and take immediate action, she will find it optimal to do so, if the value she derives given the expert's optimal learning strategy is less than that of taking any immediate action. If that is not the case, she cannot but allow the expert to continue learning.

2.7.2 Candidate payoff functions

Similarly to the discussion in Section 2.5, it will be useful to start by describing some specific learning policies. These learning polices will constitute part of the guessed equilibrium. In particular, let $V_A(p)$ and $U_A(p)$ denote the payoff functions in which $\alpha = 1$ if $p \geq \max\left\{p | V_A^E(p) = V_b(p) \bigcup U_A^P(p) = U_b(p)\right\}$ and immediate action b is taken by the principal alternatively. Lemma 2.2 implies that the principal's boundary constraint will bind first, i.e. at higher beliefs, and as a result, this policy will always be terminated by the principal. Moreover, let $V_B(p)$ and $U_B(p)$ denote the payoffs when $\alpha = 0$ if $p \leq \min\left\{p | V_B^E(p) = V_a(p) \bigcup U_B^P(p) = U_a(p)\right\}$ and immediate action a is taken by the expert alternatively. In this case the reverse will hold and the expert will be willing to take immediate action at a lower belief than the principal.

Further, we denote by $V_{cf}(p)$ and $U_{cf}(p)$, the payoffs derived when a learning policy that is identical to $U_{cf}^{E}(p)$ is pursued, i.e. to the confirmatory learning strategy that the expert pursues under complete delegation.

The final payoff function that will be relevant for deriving the optimal value functions of the game is the payoff derived by the expert, while making the principal indifferent between taking decision b immediately and allowing the expert to continue learning. In particular, the principal wants to cease learning sooner than the expert when they are sufficiently sure that the state is B, i.e., when p is sufficiently close to 0. Hence, for some region of beliefs $p \in [p', p'']$, the expert may want to persuade the principal to delay intervening as much as possible by mixing learning from information sources A and B so as to keep principal indifferent; that is, the expert chooses α to ensure that the utility derived by the principal along this interval of beliefs is

$$U(p) = U_b(p) = pu_b^A + (1-p)\,u_b^B,$$

with

$$U'(p) = u_b^A - u_b^B$$

and satisfies the HJB equation

$$rU(p) = p\alpha(p)\left(u_a^A - U(p)\right) + (1-p)(1-\alpha(p))\left(u_b^B - U(p)\right) - (2\alpha(p)-1)\,p(1-p)\,U'(p). \quad (2.16)$$

Given the linearity of $U(p)$, there exits a unique α that satisfies (2.16) and it is given by

$$\alpha^*(p) = \frac{r(1+\Delta-2p)}{p(2-\Delta)}.$$

At p', the expert sets $\alpha = 0$ and the principal is indifferent; at p'', the expert sets $\alpha = 1$ and the principal is indifferent. In between p' and p'', as the expert learns, belief p drifts down and over time the expert progressively puts more and more weight on source A; that is, chooses higher α. To see the last assertion formally, take the derivative of $\alpha^*(p)$ with respect to p

$$\alpha^{*\prime}(p) = \frac{r(\Delta+1)}{p^2(\Delta-2)}$$

Our assumption that $\rho > \Delta$ implies that the above is always negative, while the second derivative is always positive. Furthermore, there is a discontinuous jump upwards in $\alpha^*(p)$ at the belief where the principal terminates B-learning.

We call the unique learning policy $\alpha^*(p)$ *indifference learning* because its objective is to make the principal indifferent between intervening and not intervening. We denote the expert's payoff from this learning policy $V_{ind}(p)$.

2.7.3 Preliminary observations

This section develops some preliminary observations which help both in forming a guess for the flexible delegation equilibrium as well as in providing some intuition for our main result.

Lemma 2.3 shows that as the expert's bias decreases, the the r thresholds that separate the different learning strategies decrease. The lemma suggests that for the regions that learning takes place, Δ and r have a monotonic relation.

Lemma 2.3. *The threshold values of r, $\bar{r}(\Delta)$ and $\underline{r}(\Delta)$, are decreasing in Δ.*

Proof. See Appendix 2.A.14

□

Lemma 2.4. *Under complete delegation, for $r < \underline{r}(\Delta)$ and every $p \in [p^{*E}, \bar{p}^E)$*

$$U_{cf}^E(p) \geq \max\{U_a(p), U_b(p)\}.$$

Proof. See Appendix 2.A.15. □

Lemma 2.4 states that the principal will not want to terminate confirmatory learning by the expert at beliefs $p > p^{*E}$. This is an important result as it suggests that under flexible delegation, the principal would be worse off by intervening if the expert pursues his first best for $p > p^{*E}$ when $r < \underline{r}(\Delta)$. Further this lemma simplifies our analysis as if the principal did terminate confirmatory learning to the right of p^{*E}, the learning ceases to be confirmatory learning since beliefs no longer can converge to the absorbing state, as the absorbing state is outside of the learning region and the value of such learning goes down, which may make learning from the alternative source more attractive.

Lemma 2.5. $\check{p}^E \leqslant p^{*E}$

Proof. See Appendix 2.A.16. □

Lemma 2.5 is a direct consequence of the game's payoff structure and helpful towards establishing Lemma 2.6, which states for $r \in [\underline{r}(\Delta), \overline{r}(\Delta)]$ that the expert derives higher utility from learning from source B and taking immediate action a at $\overline{p}^{*E}$ than from pursuing the confirmatory strategy. This simplifies our analysis greatly, as it suggests that within this r region, the expert would never pursue the confirmatory learning policy.

Lemma 2.6. *If* $\underline{r}(\Delta) < r(\Delta) < \overline{r}(\Delta)$ then $V_B^E(p) \geqslant V_{cf}^E(p)$ *for all* p.

Proof. See Appendix 2.A.17. □

Lemma 2.7. *If* $r < \underline{r}(\Delta)$*,* $V_B^E(p) < V_{cf}^E(p)$ *for all* $p < \overline{p}^E$ *and* $V_B^E(p) > V_{cf}^E(p)$ *for all* $p > \overline{p}^E$.

Proof. See Appendix 2.A.18. □

Lemma 2.7 shows that when contradictory and confirmatory learning are both optimal under complete delegation, $V_{cf}^E(p) > V_B^E(p)$ for $p < \overline{p}^E$. Combined with Lemma 2.4 this implies that the expert will both be willing and allowed to pursue confirmatory learning under flexible delegation as he does under complete delegation.

Lemma 2.8. $U_B^E(p)$ *intersects* $U_a(p)$ *from above, is strictly convex and*

$$U_B^E(p) > \max\{U_a(p), U_b(p)\} \text{ for } p \in [\check{p}^E, \overline{p}^{*E}).$$

Proof. See Appendix 2.A.19.

□

Lastly, Lemma 2.8 is pivotal in establishing that the principal is better off when the expert pursues the latter's optimal contradictory strategy, than taking her outside option for $p > \overline{p}^{*E}$. Combined with the previous lemmas, this means that for $p > \overline{p}^{*E}$ the expert can pursue his first-best knowing that the principal would not intervene.

2.7.4 Equilibrium under Flexible Delegation

In this section we provide a description of the equilibrium under flexible delegation, a sketch of the proof as well as the main insights of equilibrium. A detailed proof of the equilibrium characterisation can be found in Appendix 2.A.23.

Firstly, Corollary 2.1 shows that for different combinations of r and Δ the optimal learning strategy may be either contradictory only or contradictory and confirmatory. Further, Lemma 2.9 establishes the regions in which the expert is better off by using a policy that keeps the principal indifferent between $Y = b$ and $Y = \emptyset$. Given these results, the equilibrium we proceed by guessing the equilibrium of flexible delegation.

To confirm that this policy constitutes an equilibrium, one needs to show that for each of these cases the below both the HJB's for the principal and the expert are satisfied. That is, one needs to show that:

$$\max\left\{\max_{\alpha} H^E(\alpha) - rV(p),\ V^E(p) - V(p)\right\} = 0 \tag{2.17}$$

$$\max\left\{H^P(\alpha^*) - rU(p),\ U^P(p) - U(p)\right\} = 0 \tag{2.18}$$

with

$$H^E(\alpha) \equiv p\alpha(v_a^A - V(p)) + (1-p)(1-\alpha)(v_b^B - V(p)) - (2\alpha - 1)p(1-p)V'(p)$$

and

$$H^P(\alpha) \equiv p\alpha^*(u_a^A - U(p)) + (1-p)(1-\alpha^*)(u_b^B - U(p)) - (2\alpha^* - 1)p(1-p)U'(p)$$

Therefore, to characterise the equilibrium we guess a policy and ensure that the above conditions hold.[16] The extended verification of the below equilibrium can be found in Appendix 2.A.23.

The equilibrium of the game is as follows:

- If $\bar{r}(\Delta) < r(\Delta)$ then the principal's boundary condition 2.8 binds and $y =$

[16] We check the optimality of the conjectured learning policy at all points of differentiability of the expert and the principal's value functions. However, under flexible delegation the value functions have points of non-differentiability and so further work has to be done to verify that the derived value is a viscosity solution.

argmax $U_x(p)$.

- If $\bar{r}(\Delta) > r(\Delta) > \underline{r}(\Delta)$ and $V_A\left(\underline{p}^{*P}\right) \geq V_B\left(\underline{p}^{*P}\right)$

$$\begin{cases} y = b & \text{for } p \in \left[0, \underline{p}^{*P}\right] \\ y = \emptyset, \alpha = 1 & \text{for } p \in \left(\underline{p}^{*P}, \check{p}^{E}\right) \\ y = \emptyset, \alpha = 0 & \text{for } p \in \left(\check{p}^{E}, \overline{p}^{*E}\right) \\ x = a & \text{for } p \in \left[\overline{p}^{*E}, 1\right] \end{cases}$$

- If $\bar{r}(\Delta) > r(\Delta) > \underline{r}(\Delta)$ and $V_A\left(\underline{p}^{*P}\right) < V_B\left(\underline{p}^{*P}\right)$

$$\begin{cases} y = b & \text{for } p \in \left[0, \underline{p}^{*P}\right] \\ y = \emptyset, \alpha = \alpha^*(p) & \text{for } p \in \left(\underline{p}^{*P}, \widehat{p}\right) \\ y = \emptyset, \alpha = 1 & \text{for } p \in \left[\widehat{p}, \check{p}^{E}\right) \\ y = \emptyset, \alpha = 0 & \text{for } p \in \left(\check{p}^{E}, \overline{p}^{*E}\right) \\ x = a & \text{for } p \in \left[\overline{p}^{*E}, 1\right] \end{cases}$$

- If $r(\Delta) < \underline{r}(\Delta)$ and $V_A\left(\underline{p}^{*P}\right) \geq V_{cf}\left(\underline{p}^{*P}\right)$

$$\begin{cases} y = b & \text{for } p \in \left[0, \underline{p}^{*P}\right] \\ y = \emptyset, \alpha = 1 & \text{for } p \in \left(\underline{p}^{*P}, p''\right) \\ y = \emptyset, \alpha = 0 & \text{for } p \in \left[p'', p^{*E}\right) \\ y = \emptyset, \alpha = 1/2 & \text{for } p = p^{*E} \\ y = \emptyset, \alpha = 1 & \text{for } p \in \left(p^{*E}, \overline{p}^{E}\right) \\ y = \emptyset, \alpha = 0 & \text{for } p \in \left[\overline{p}^{E}, \overline{p}^{*E}\right) \\ x = a & \text{for } p \in \left[\overline{p}^{*E}, 1\right] \end{cases}$$

- If $r(\Delta) < \underline{r}(\Delta)$ and $V_A\left(\underline{p}^{*P}\right) < V_{cf}\left(\underline{p}^{*P}\right)$

$$\begin{cases} y = b & \text{for } p \in \left[0, \underline{p}^{*P}\right] \\ y = \emptyset, \alpha = \alpha^*(p) & \text{for } p \in \left(\underline{p}^{*P}, \widetilde{p}\right) \\ y = \emptyset, \alpha = 0 & \text{for } p \in \left[\widetilde{p}, p^{*E}\right) \\ y = \emptyset, \alpha = 1/2 & \text{for } p = p^{*E} \\ y = \emptyset, \alpha = 1 & \text{for } p \in \left(p^{*E}, \overline{p}^{E}\right) \\ y = \emptyset, \alpha = 0 & \text{for } p \in \left[\overline{p}^{E}, \overline{p}^{*E}\right) \\ x = a & \text{for } p \in \left[\overline{p}^{*E}, 1\right] \end{cases}$$

The equilibrium is comprised of many different cases and sub-cases, which makes it difficult to form an intuitive understanding of optimal play. However, comparing it to complete delegation in which the expert learns about the state optimally, without the principal's intervention, can help highlight the interesting features of the equilibrium. In particular, the first thing to note is that in the case of contradictory learning the principal never intervenes and the expert always pursues his first best for any $p > \check{p}^E$. The same holds in the case of both contradictory and confirmatory learning for any $p > \max\{p'', \tilde{p}\}$. In both these regions, complete and flexible delegation are payoff equivalent for both the principal and the expert and the optimal policy is equivalent to that of the expert's optimal learning.

Another feature of the equilibrium is that, wherever necessary, for $p > \underline{p}^{*P}$ the expert is always better of pursuing a policy that makes the principal indifferent than taking immediate action. This is shown in Lemma 2.9 below and implies that immediate action will never be taken for $p \in [\underline{p}^{*P}, \overline{p}^{*E}]$ for $r < \bar{r}(\Delta)$.

Lemma 2.9. *$V_{ind}(p) \geq V_b(p)$ whenever the optimal learning policy in the decision problem, α, is 1 and the immediate action optimal decision is b.*

Proof. See Appendix 2.A.21. ☐

Lemma 2.9 states that the expert derives a higher payoff from making the principal indifferent than from taking decision b, in the regions where the optimal policy for the expert in the decision problem is to learn with intensity $\alpha = 1$.

The third notable feature of the equilibrium has to do with the switching of optimal strategies and explains our Theorem 2.1. In particular if $r(\Delta) < \underline{r}(\Delta)$ and $V_A\left(\underline{p}^{*P}\right) < V_{cf}\left(\underline{p}^{*P}\right)$ the expert pursues confirmatory learning after $\tilde{p}$ with $\alpha = 0$. This differs from the experts optimal learning under complete delegation in this region where he would pursue contradictory learning with $\alpha = 1$. The intuition behind this learning shift is discussed in the next section and is the main contribution of our paper.

Further, Corollary 2.1 below, shows that given our assumptions all of the discussed above are non-empty. Specifically, Figure 2.7 shows that the threshold r below which both confirmatory and contradictory learning is optimal. As can be seen from the figure, there exists a threshold for r below which both confirmatory and contradictory strategies must be part of the equilibrium as the highest Δ dictated by r^* is lower than the lowest Δ required for contradictory strategy to be optimal only.

Corollary 2.1. *There exists a threshold for r, below which both confirmatory and contradictory strategies are pursued in equilibrium, for some p.*

Proof. See Appendix 2.A.22.

□

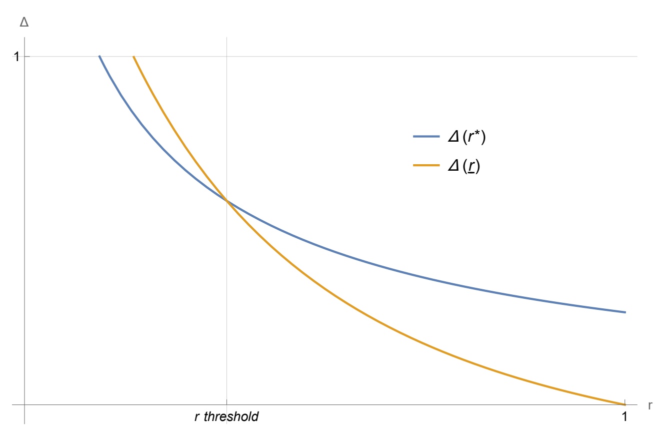


Figure 2.7: Learning strategy regions

2.8 Comparison of Complete and Flexible Delegation

Theorem 2.1 states that complete delegation may be detrimental to the principal when the principal's and the expert's preferences are sufficiently closely aligned. The result is intuitive for both types of the expert's optimal learning policies and is driven by the expert's strategic response to the principal's intervention. In this section, we discuss the intuition supporting Theorem 2.1.[17]

Under flexible delegation, the ability to override the expert enables the principal to stop learning as soon as the posterior belief falls sufficiently to make immediate decision optimal. This is the main benefit of flexible delegation. However, the possibility of the principal terminating learning triggers the expert's strategic response. This strategic response may hurt the principal. Theorem 2.1 charaterises the conditions under which the expert's strategic response makes the principal worse off.

In the case of only contradictory learning, if $U^E\left(\check{p}^E\right) \leq U_b\left(\check{p}^E\right)$ – that is, if at the belief where the autonomous expert switches from contradictory learning from source A to contradictory learning from B, the principal's payoff under expert's learning is lower than the payoff from immediate decision, – an intervention by the principal does not trigger a strategic response from the expert. Hence, flexible delegation does not have a cost; it unambiguously improves the principal's payoff. Similarly, for the case of mixed learning, if $U^E\left(\underline{p}^{*E}\right) \leq U_b\left(\underline{p}^{*E}\right)$ – that is, if at the belief where the autonomous expert switches from contradictory A to confirmatory B learning, the principal's payoff under expert's learning is lower than the payoff from immediate decision,— an intervention by the principal does not trigger a strategic response from the expert.

However, if $U^E\left(\check{p}^E\right) > U_b\left(\check{p}^E\right)$ (in the case of contradictory learning) or $U^E\left(\underline{p}^{*E}\right) > U_b\left(\underline{p}^{*E}\right)$ (in the case of mixed learning), an intervention by the principal changes the expert's optimal learning strategy. In particular, for sufficiently small Δ, some contradictory learning from source A remains optimal for low beliefs, but the expert switches to A learning later than he would have done absent the principal's intervention. The delay in the switch to A learning hurts the principal as by Lemma 2.2, the principal wants the expert to switch to A learning early. On the other hand, in this region under complete delegation the expert may want to pursue A learning instead, as he would not need to act strategically, which would make the principal better off.

Figure 2.8 demonstrates this. Under flexible delegation, the optimal learning policy by the expert will keep the principal's utility equal to $U_b(p)$ for $p \in [\underline{p}^{*P},\ p_{ind}]$. However, for some $p > \underline{p}^{*P}$ the principal would be better off if the expert pursued his

[17] For formal proof see Appendix 2.A.24.

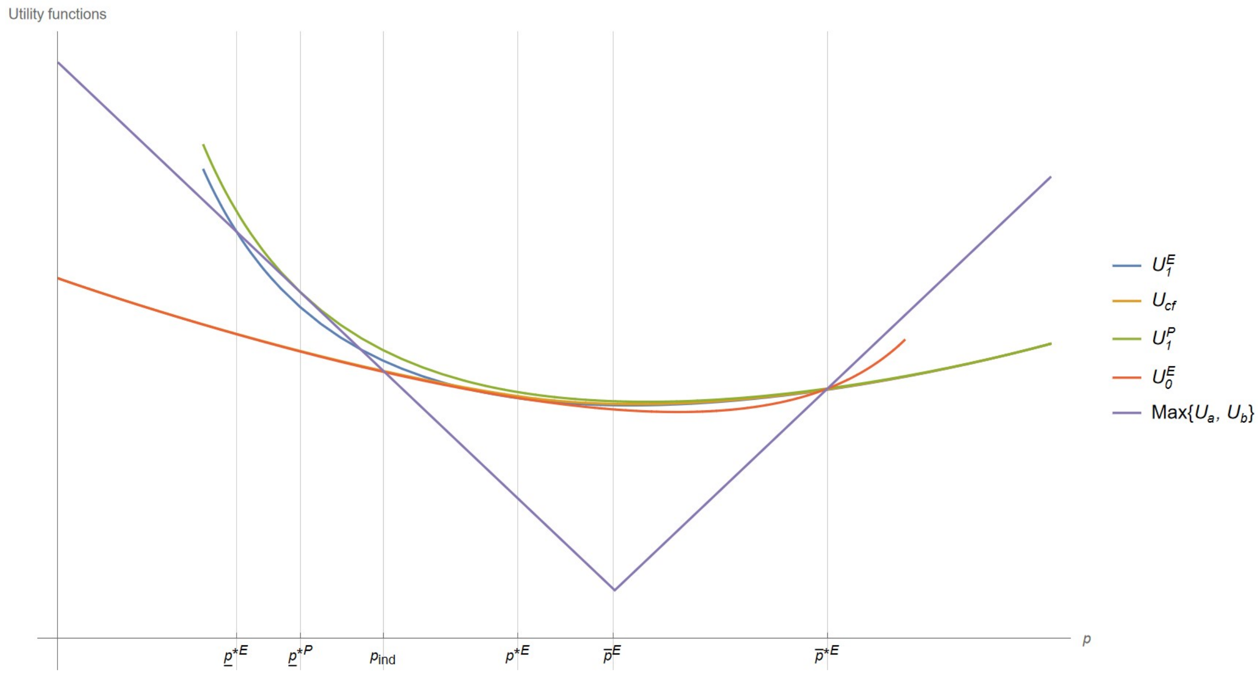


Figure 2.8: Example of complete delegation being optimal for the principal

first-best learning policy which for $p \in [\underline{p}^{*E},\ p^{*E}]$ is learning from source A (Blue line).

2.9 Concluding Remarks

We study a dynamic model in which the principal delegates learning about the binary state of the world to a biased expert who has access to a state confirming learning technology. The principal can either delegate learning and the decision making to the expert or can retain the ability to make decisions while the expert is still learning. We call the former complete, and the latter flexible, delegation.

We show that when the expert's preferences are sufficiently different from the principal's preferences, the principal is always better off under flexible delegation. However, for low level of expert's bias, there exist beliefs about the state for which the principal is better off under complete delegation.

To prove our main result, we fully characterize expert's equilibrium learning strategy under flexible delegation. We find that on the equilibrium path, from the perspective of the principal three types of learning sub-optimalities may occur. Firstly, the expert may terminate learning at lower beliefs than those that the principal would prefer. As the expert has the ability to make a decision, the principal cannot affect this outcome. Secondly, for some beliefs, the expert still wants to learn, while the principal wants to take a decision. Thirdly, for some beliefs, the principal and the expert may disagree on the information source from which to learn. The possibility of principal's intervention eliminates the cost of expert's prolonged learning, but makes expert learning strategy even more sub-optimal than it is under complete delegation.

Appendix

Appendix 2.A

2.A.1 Evolution of posterior beliefs

Note that if the signal has not arrived in the time interval of length h, then by Bayes' rule,

$$
\begin{aligned}
p_{t+h} &= \mathbb{P}\left(\omega = A \mid \text{no signal}\right) \\
&= \frac{\mathbb{P}\left(\text{no signal} \mid \omega = A\right)\mathbb{P}\left(\omega = A\right)}{\mathbb{P}\left(\text{no signal}\right)} \\
&= \frac{\left(e^{-\alpha h} \times 1\right) p_t}{\left(e^{-\alpha h} \times 1\right) p_t + \left(1 \times e^{-(1-\alpha)h}\right)\left(1 - p_t\right)} \\
&= \frac{p_t e^{-\alpha h}}{p_t e^{-\alpha h} + \left(1 - p_t\right) e^{-(1-\alpha)h}} \\
&= p_t - \left(2\alpha - 1\right) p_t \left(1 - p_t\right) h,
\end{aligned}
$$

where the last equality follows by the Taylor approximation.

2.A.2 Derivation of $U^*(p)$

$$\begin{aligned}
U^*(p) &= \mathbb{E}\left[e^{-r\tau}u_a^A \mid \omega = A\right] + \mathbb{E}\left[e^{-r\tau}u_b^B \mid \omega = B\right] \\
&= pu_a^A\mathbb{E}\left[e^{-r\tau}\right] + (1-p)u_b^B\mathbb{E}\left[e^{-r\tau}\right] \\
&= \left(pu_a^A + (1-p)u_b^B\right)\mathbb{E}\left[e^{-r\tau}\right] \\
&= \left(pu_a^A + (1-p)\,u_b^B\right)\int_0^\infty e^{-rt}\left[\frac{e^{-t/2}}{2}\right]dt \\
&= \left(pu_a^A + (1-p)\,u_b^B\right)\frac{1}{2r+1} \\
&= \frac{\Delta(1-p)+1}{2r+1}
\end{aligned}$$

where, $\tau = \min\left\{\tau^a, \tau^b\right\}$, the time that the first, and only, signal arrives.

2.A.3 Derivation of $\underline{p}^{*p}$

By definition, $\underline{p}^{*p} \equiv \{p \mid U_A^P(p) = U_b(p)\}$. That is, it is the belief that principal's utility from immediate actions is the same as the expected utility from learning from source A and acting only when signal A arrives.

$$U_b(p) = p\ u_b^A + (1-p)\ u_b^B = \Delta + 1 - 2p \tag{2.19}$$

The principal's HJB, equation 2.13, when $\alpha = 1$ is

$$\begin{aligned}
rU^P(p) &= \alpha p\left(u_a^A - U^P(p)\right) + (1-\alpha)(1-p)\left(u_b^B - U^P(p)\right) - (2\alpha-1)\,p\,(1-p)\,U^{P\prime}(p) \\
rU_A^P(p) &= p\left(u_a^A - U_A^P(p)\right) - p\,(1-p)\,U_A^{P\prime}(p)
\end{aligned}$$

where subscript A denotes the HJB when $\alpha = 1$. Smooth pasting implies that $U_b'(\underline{p}^{*p}) = U_A'^P(\underline{p}^{*p}) = -2$. Hence,

$$U_A^P(\underline{p}^{*p}) = \frac{(\underline{p}^{*p})^2(-2) + 3\underline{p}^{*p}}{\underline{p}^{*p} + r} \tag{2.20}$$

Setting (2.20) = (2.19) and solving for $\underline{p}^{*p}$

$$\underline{p}^{*p} = \frac{r(\Delta + 1)}{2(r+1) - \Delta}$$

2.A.4 Derivation of contradictory learning

Focusing on source A , i.e. $\alpha = 1$, reduces (2.9) to

$$rU_A^P(p) = p\left(u_a^A - U_A^P(p)\right) - p\,(1-p)\,U_A'(p)$$

given the boundary condition (2.14). This ODE is readily solvable and applying our assumptions is equal to

$$U_A^P(p) = \begin{cases} \frac{p+(1-p)(\Delta+1)\left(\frac{(p-1)r(\Delta+1)}{\Delta p(r+1)-p(r+2)}\right)^r}{r+1} & p \geq \underline{p}^{*P} \\ U_b\,(p) & p < \underline{p}^{*P} \end{cases}$$

where $\underline{p}^{*P}$is the belief p' that satisfies the condition $U_A^P(p') = U_b(p')$.[18] As $\alpha = 1$, conditional on no signal arriving, Bayesian updating dictates that beliefs are going to move away from state A. Intuitively, as attention is on source A, since no signal arrives, the principal becomes more pessimistic about the state being A.

Similarly, learning from source B only, the principal derives the following payoff:

$$U_B^P(p) = \begin{cases} \frac{p(pr)^r(-((p-1)(\Delta+(r+2))))^{-r}-(p-1)(\Delta+1)}{r+1} & p \leq \overline{p}^{*P} \\ U_a\,(p) & p > \overline{p}^{*P} \end{cases}$$

where $\overline{p}^{*P}$ is belief p that satisfies the condition $U_B^P(p;\cdot) = U_a(p;\cdot)$.[19]

2.A.5 Derivation of $\overline{p}^{*P}$

By definition, $\overline{p}^{*P}$ satisfies, $U_B^P(\overline{p}^{*P};\cdot) = U_a(\overline{p}^{*P};\cdot)$.
Hence,

$$U_a(p) = p\,U_A^a + (1-p)\,U_B^a = 2p - 1 \tag{2.21}$$

[18]See Appendix 2.A.3
[19]See Appendix 2.A.5

The principal's HJB, equation 2.13, when $\alpha = 0$ is

$$rU^P(p) = \alpha p\left(u_a^A - U^P(p)\right) + (1-\alpha)(1-p)\left(u_b^B - U^P(p)\right) - (2\alpha - 1)\,p(1-p)\,U^{P\prime}(p)$$
$$rU_B^P(p) = (1-p)\left(u_b^B - U_B^P(p)\right) + p(1-p)\,U_B^{P\prime}(p)$$

where subscript B denotes the HJB when $\alpha = 0$. Smooth pasting implies that $U_a'(\overline{p}^{*p}) = U_B'^P(\overline{p}^{*p}) = 2$. Hence,

$$U_B^P(\overline{p}^{*p}) = \frac{(\overline{p}^{*p} - 1)(\Delta + 2\overline{p}^{*p} + 1)}{\overline{p}^{*p} - r - 1} \tag{2.22}$$

Setting (2.22) = (2.21) and solving for $\overline{p}^{*p}$

$$\overline{p}^{*p} = \frac{\Delta + (r+2)}{\Delta + 2(r+1)}$$

2.A.6 Derivation of p^{*P}

Recall that

$$U^*(p) = \frac{\Delta(1-p) + 1}{2r+1}$$

By smooth pasting, as $p \to p^{*P}$ from above, i.e. when $\alpha = 1$, the HJB equation 2.9 must equal $U^*(p)$.

Then,

$$rU^P(p^{*P}) = \alpha p^{*P}\left(u_a^A - U^P(p^*)\right) + (1-\alpha)(1-p^*)\left(u_b^B - U^P(p^*)\right) - (2\alpha-1)\,p^{*P}\left(1-p^{*P}\right)U^{P\prime}(p^{*P})$$
$$rU^*(p^{*P}) = p^{*P}\left(u_a^A - U^*(p^{*P})\right) - p^{*P}\left(1-p^{*P}\right)U'(p^{*P})$$
$$r\left(p^{*P}u_a^A + \left(1-p^{*P}\right)u_b^B\right)\frac{1}{2r+1} = p^{*P}\left(u_a^A - \left(p^{*P}u_a^A + \left(1-p^{*P}\right)u_b^B\right)\frac{1}{2r+1}\right) - p^{*P}\left(1-p^{*P}\right)\frac{u_a^A - u_b^B}{2r+1}$$
$$p^{*P} = \frac{\upsilon_b^B}{\upsilon_b^B + \upsilon_a^A} = \frac{1+\Delta}{2+\Delta}$$

2.A.7 Derivation of confirmatory learning

Postulating smooth pasting as $p \to p^{*P}$ both from below and above, as it must be the case that payoffs match, since strategies at belief p^{*P} are identical and applying these in (2.9) we get

$$p^{*P}=\frac{1+\Delta}{2+\Delta} \quad \text{[20]} \tag{2.23}$$

Clearly, at belief p^{*P}, the principal's utility is equal to $U^{*}(p^{*P})$, as derived above. Given this boundary condition we can solve 2.9. In addition, note that this learning policy results in the principal always taking the decision that matches the state, as she always acts after the arrival of a signal. The payoff of the confirmatory learning strategy , hence forth to be denoted by U_{cf}^{P}, is

$$U_{cf}^{P}(p)=\begin{cases}\frac{p\left(-\frac{p}{(p-1)(\Delta+1)}\right)^{r}-(p-1)(2r+1)(\Delta+1)}{(r+1)(2r+1)} & p<p^{*P}\\ \frac{\Delta(1-p)+1}{2r+1} & p=p^{*P}\\ \frac{p\left(-\frac{(p-1)(\Delta+1)}{p}\right)^{r+1}+p(2r+1)}{(r+1)(2r+1)} & p>p^{*P}\end{cases}$$

2.A.8 Derivation of $\bar{r}^{P}(\Delta)$

The threshold $\bar{r}^{P}(\Delta)$ is the smallest discount rate r' such that the principal prefers choosing the action immediately to learning for $\delta\rightarrow 0$ units of time and only then choosing the optimal action. To find $\bar{r}^{P}(\Delta)$: Suppose $p<\hat{p}^{P}$ and recall that $U_b(p)=pu_b^A+(1-p)\,u_b^B$. If the principal learns from source A for δ amount of time and then chooses the optimal action, she gets:

$$\begin{aligned}e^{-r\delta}\left[\left(1-e^{-\delta}\right)pu_a^A+\left(1-p+e^{-\delta}p\right)U_b\left(p'\right)\right]\approx\\ U_b(p)-rU_b(p)\,\delta+\underbrace{\left(u_a^A-U_b(p)\right)}_{=u_a^A-u_b^A+\left(u_b^A-u_b^B\right)(1-p)}p\delta-\underbrace{U_b'(p)}_{=u_b^A-u_b^B}p(1-p)\,\delta=\\ U_b(p)-r\left(u_b^B+\left(u_b^A-u_b^B\right)p\right)\delta+\left(u_a^A-u_b^A\right)p\delta\end{aligned}$$

Learning is not optimal if

$$e^{-r\delta}\left[\left(1-e^{-\delta}\right)pu_a^A+\left(1-p+e^{-\delta}p\right)U_b\left(p'\right)\right]\leq U_b(p)$$

[20]See Appendix 2.A.6

for every $p < \hat{p}^P$. That is,

$$-ru_b^B + r\left(u_b^B - u_b^A\right)p + \left(u_a^A - u_b^A\right)p \leq 0.$$

The LHS is increasing in p. So set

$$p = \hat{p}^P = \frac{u_b^B - u_a^B}{u_b^B - u_a^B + u_a^A - u_b^A}$$

Then

$$r \geq \frac{\left(u_b^B - u_a^B\right)\left(u_a^A - u_b^A\right)}{u_b^B u_a^A - u_a^B u_b^A}$$

The argument for $p > \hat{p}^P$ is symmetric and should yield the same threshold $\bar{r}^P(\Delta)$. Substituting our assumptions on u_x^ω yields:

$$r \geq \frac{4 - \Delta^2}{2\Delta} \equiv \bar{r}^P(\Delta).$$

2.A.9 Derivation of $\underline{r}^P(\Delta)$

In order to characterize the threshold level $\underline{r}^P(\Delta)$ we make the following observation.

Lemma 2.10. *$U_A^P(p^{*P}) > U_B^P(p^{*P})$ for $\Delta > 0$.*

Proof. Condition

$$U_A^P(p^{*P}; \cdot) > U_B^P(p^{*P}; \cdot)$$

can be equivalently expressed as

$$\frac{(\Delta+1)\left(\left(\frac{r}{(r+2)-\Delta(r+1)}\right)^r + 1\right)}{(r+1)(\Delta+2)} > \frac{(\Delta+1)(\Delta+(r+2))^{-r}(r(\Delta+1))^r + (\Delta+(r+2))^r}{(r+1)(\Delta+2)}$$

After some algebraic manipulations this becomes

$$\left(\frac{r}{(r+2)-\Delta(r+1)}\right)^r > \left(\frac{r(\Delta+1)}{\Delta+(r+2)}\right)^r$$

or

$$-\frac{\Delta^2 r(r+1)}{(\Delta(r+1)-(r+2))(\Delta+(r+2))} > 0$$

The above is satisfied for $1 > \Delta > 0$.

$\square$

Further, we use Proposition 6, Che and Mierendorff (2019), which states that the confirmatory strategy is part of the optimal strategy if and only if at p^{*P} the principal finds it optimal to pursue the confirmatory strategy. Hence, the confirmatory strategy will not be part of the optimal strategy for the principal, iff:

$$U_A^P(p^{*P}) \geq U^*(p^{*P})$$

$$\frac{(\Delta+1)\left(\left(\frac{r}{-\Delta(r+1)+r+2}\right)^r+1\right)}{(\Delta+2)(r+1)} \geq \frac{2(\Delta+1)}{(2r+1)(\Delta+2)}$$

$$\left(\frac{r}{r+2-\Delta(r+1)}\right)^r \geq \frac{1}{2r+1}$$

The value of r that makes the above condition hold with equality is the threshold value $\underline{r}^P(\Delta)$. For a given r, the highest level of Δ for which the optimal strategy includes a confirmatory and contradictory part is:

$$\Delta = \frac{-r(2r+1)^{\frac{1}{r}}+r+2}{r+1} \tag{2.24}$$

2.A.10 Derivation of p^{*E}

Recall that

$$V^*(p) = \frac{\Delta p + 1}{2r+1}$$

By smooth pasting, as $p \to p^{*E}$ from above, i.e. when $\alpha = 1$. The HJB for the expert must satisfy $V^*(p)$

Then,

$$rV^E(p^{*E}) = \alpha p^{*E}\left(v_a^A - V^E(p^*)\right) + (1-\alpha)(1-p^*)\left(v_b^B - V^E(p^*)\right) - (2\alpha-1)p^{*E}\left(1-p^{*E}\right)V^{E\prime}(p^{*E})$$
$$rV^*(p^{*E}) = p^{*E}\left(v_a^A - V^*(p^{*E})\right) - p^{*E}\left(1-p^{*E}\right)V'(p^{*E})$$
$$r\left(p^{*E}v_a^A + \left(1-p^{*E}\right)v_b^B\right)\frac{1}{2r+1} = p^{*E}\left(v_a^A - \left(p^{*E}v_a^A + \left(1-p^{*E}\right)v_b^B\right)\frac{1}{2r+1}\right) - p^{*E}\left(1-p^{*E}\right)\frac{v_a^A - v_b^B}{2r+1}$$
$$p^{*E} = \frac{v_b^B}{v_b^B + v_a^A} = \frac{1}{\Delta+2}$$

2.A.11 Proof of Lemma 2.2

Lemma 2.2. For any $\Delta > 0$, the expert's learning region is shifted to the left relative to the principal's learning region. Specifically, $\underline{p}^{*E} < \underline{p}^{*P}$, $\bar{p}^{*E} < \bar{p}^{*P}$ and $p^{*E} < p^{*P}$.

Proof. As we have analytical solutions for all thresholds the proof can be produced by comparing the learning regions.

$\underline{\mathbf{p}}^{*E} < \underline{\mathbf{p}}^{*P}$

$$\underline{p}^{*P} - \underline{p}^{*E} = \frac{\Delta r(\Delta + 4 + 2r)}{(2(1+r) - \Delta)(\Delta + 2(1+r))},$$

which is greater than zero since $1 > \Delta$ by assumption and $r \geq 0$.

$\bar{p}^{*E} < \bar{p}^{*P}$

$\bar{p}^{*P} - \bar{p}^{*E} = -\dfrac{\Delta r(\Delta + 2(r+2))}{(\Delta - 2(1+r))(\Delta + 2(1+r))}$, which is greater than zero if $\Delta < 2(1+r)$.

This is clearly the case since $1 > \Delta$ and $r > 0$.

$p^{*E} < p^{*P}$

$p^{*P} - p^{*E} = \frac{1+\Delta}{\Delta+2} - \frac{1}{\Delta+2} = \frac{\Delta}{\Delta+2} > 0$ for $\Delta > 0$

□

2.A.12 Derivation of Restriction 2.1

We impose the restriction that if the expert find's it optimal to learn forever, then the principal also prefers learning to her outside option, i.e.

$$U^E_{cf}(p^{*E}) - U_b\left(p^{*E}\right) = \frac{2 - \Delta - 2(\Delta + 3)\Delta r}{(\Delta + 2)(2r + 1)} > 0 \tag{2.25}$$

For $r = 0$, the above reduces to

$$\frac{2 - \Delta}{\Delta + 2},$$

which is positive. Further, the derivative of (2.25) with respect to r is:

$$-\frac{2(\Delta^2+2\Delta+2)}{(\Delta+2)(2r+1)^2}<0,$$

which implies that as r increases, $U_{cf}^{E}\left(p^{*E}\right)-U_b\left(p^{*E}\right)$ decreases monotonically. Setting $U_{cf}^{E}\left(p^{*E}\right)-U_b\left(p^{*E}\right)=0$ and solving for r^* yields

$$r^*=\frac{2-\Delta}{2\Delta(\Delta+3)}.$$

Hence, the required restriction is

Restriction 2.2.

$$r<\frac{2-\Delta}{2\Delta(\Delta+3)}\equiv r^*$$

Equivalently, we can express Restriction 2.2 as a restriction on Δ. In particular, solving for Δ yields:

$$\bar{\Delta}\left(r\right)=\frac{\sqrt{36r^2+28r+1}-6r-1}{4r}>0 \tag{2.26}$$

Therefore, $\Delta<\max\{\bar{\Delta}\left(r\right),1\}$ is our Restriction 2.1.

2.A.13 Implication of Restriction 2.2

In order to agree on the optimal decision when the expert stops learning it must be the case that

$$\bar{p}^{*E}\geq\hat{p}^{P}$$

Indeed,

$$\frac{r(\Delta+1)}{\Delta-2(r+1)}+1\geq\frac{\Delta+2}{4} \tag{2.27}$$

where $\hat{p}^P$ is such that $U_a\left(\hat{p}^P\right)=U_b\left(\hat{p}^P\right)$, that is,

$$\hat{p}^P=\frac{\Delta+2}{4}$$

.

(2.27) reduces to

$$r \leq \frac{(2-\Delta)^2}{6\Delta}$$

This condition is always met as restriction 2.2 requires;

$$r^* < \frac{2-\Delta}{2\Delta(\Delta+3)} < \frac{(2-\Delta)^2}{6\Delta} \tag{2.28}$$

2.A.14 Proof of Lemma 2.3

Proof.

$$\bar{r}\,(\Delta) = \frac{4-\Delta^2}{2\Delta}$$
$$\frac{d}{d\Delta}[\bar{r}\,(\Delta)] = -\frac{\Delta^2+4}{2\Delta^2}$$

which is clearly negative for any $\Delta \neq 0$.

Further, the condition $\underline{r}\,(\Delta)$ must satisfy, if it exists, is

$$\Delta = \frac{\left(-r(2r+1)^{\frac{1}{r}} + r + 2\right)}{r+1} \tag{2.29}$$

Let,

$$0 = \Delta - \frac{\left(-r(2r+1)^{\frac{1}{r}} + r + 2\right)}{r+1}$$
$$0 = \frac{\partial}{\partial\Delta} d\Delta + \frac{\partial}{\partial r} dr$$

Then,

$$\frac{dr}{d\Delta} = \frac{r(r+1)^2(2r+1)}{2r^2 + r + (2r+1)^{\frac{1}{r}}(r(4r+3) - (r+1)(2r+1)\log(2r+1))}$$

given that $r > 0$ the above expression is negative if

$$(2r+1)^{\frac{1}{r}}((r+1)(2r+1)\log(2r+1)-r(4r+3)) < r(2r+1)$$
$$\frac{(2r+1)^{\frac{1}{r}-1}((r+1)(2r+1)\log(2r+1)-r(4r+3))}{r} < 1$$

The above inequality holds for any r between 0 and 3.5271.

Finally note that if $r > 1$, (2.29) is negative which cannot be as Δ must be positive. Hence, the upper bound of $\underline{r}(\Delta)$ is 1, which concludes the proof.

□

2.A.15 Proof of Lemma 2.4

Lemma 2.4. Under complete delegation, for every $p > p^{*E}$

$$U_{cf}^{E}(p) \geq \max\{U_a(p), U_b(p)\}.$$

Proof. First, it is immediate that $U_b\left(p^{*E}\right) > U_a\left(p^{*E}\right)$. Indeed,

$$U_a\left(p^{*E}\right) = -\frac{\Delta}{\Delta+2} < 0$$

and

$$U_b\left(p^{*E}\right) = \frac{\Delta(\Delta+3)}{\Delta+2} > 0$$

Then, note that $U_{cf}^{E}\left(p^{*E}\right) > U_b\left(p^{*E}\right)$ by restriction 2.2.

Note also that $U_{cf}^{E'}(p^{*E})$ with respect to r is:

$$-\frac{\Delta\left(\Delta+(\Delta+2)^2 r+1\right)}{(\Delta+1)(2r+1)} < 0$$

since, $U_b'(p) = -2$, it is always the case that $U_{cf}^{E'}(p^{*E}) > U_b'(p^{*E})$ for $r < r^*$.

It remains to show that if $U_{cf}^{E}(p') < \max\{U_a(p'), U_b(p')\}$ for some $p' > p^{*E}$, $U_{cf}^{E}(p')$ would never be pursued by the expert in equilibrium. To do so, recall that the the utility derived by the principal from the experts learning decreases monotonically in r and Δ. Hence the limiting case as $r \to r^*$ and $\Delta \to 1$ is the lower bound of $U_{cf}^{E'}(p)$ conditional on p.

Then,

$$U_{cf}^{E}(p) - U_a(p)|_{r=r^*,\ \Delta=1} = \frac{-10p^{9/8} + 7\ 2^{7/8}(1-p)^{9/8} + 9\sqrt[8]{p}}{9\sqrt[8]{p}}$$

which is equal to zero for $p' = 0.947192$. However, $\overline{p}^{*E}|_{r=r^*,\ \Delta=1} = 0.8$. This means that the expert would always take action a for $p > \overline{p}^{*E}$ and hence $U_{cf}^{E}(p')$ will never be part of the experts equilibrium learning. Finally, note that $U_a(p') = 0.8944 > 0.1056 = U_b(p')$, which concludes the proof for $r < r^*$.

□

2.A.16 Proof of Lemma 2.5

Proof. Simple algebra shows that $V_B^E\left(p^{*E}\right) > V_A^E\left(p^{*E}\right)$ for $\Delta > 0$. Hence, the expert always prefers learning from source B at p^{*E}. Since $V_B^E(p)$ and $V_A^E(p)$ only cross once and learning from source A is always preferred for low enough p, while learning from source B is always preferred for high enough p, it follows that for any $p \geq p^{*E}$learning from source B is always preferred by the expert and thus $\check{p}^E \leqslant p^{*E}$.

$$V_B^E\left(p^{*E}\right) > V_A^E\left(p^{*E}\right)$$

$$\frac{(\Delta+1)((r+2)-\Delta(r+1))^{-r}\left(((r+2)-\Delta(r+1))^r + (r)^r\right)}{(r+1)(\Delta+2)} > \frac{(\Delta+1)\left(\frac{1}{\Delta+2}\right)^{1-r}\left(\left(\frac{1}{\Delta+2}\right)^r + \left(\frac{r(\Delta+1)}{(\Delta+2)(\Delta+(r+2))}\right)^r\right)}{r+1}$$

After some algebraic manipulation the above inequality reduces to

$$-\frac{\Delta^2 r(r+1)(\Delta+1)}{(\Delta(1+r)-(2+r))(\Delta+2+r)} > 0$$

which is always met for $1 > \Delta > 0$ and $r > 0$.

□

2.A.17 Proof of Lemma 2.6

Proof. Since, $\underline{r}(\Delta) < r(\Delta) < \overline{r}(\Delta)$, only contradictory strategy would be used by the expert under complete delegation for any p. Hence, $V_B^E\left(\check{p}^E\right) = V_A^E\left(\check{p}^E\right) \geqslant V_{cf}^E\left(\check{p}^E\right)$. In addition, $V_B^E(p) \geqslant V_{cf}^E(p)$ for any belief $p \geqslant \check{p}^E$, since contradictory from source B is the expert's optimal learning strategy under complete delegation for these beliefs. Also, from Lemma 2.5 we know that $\check{p}^E \leqslant p^{*E}$, which implies that $V_B^E\left(p^{*E}\right) > V_{cf}^E\left(p^{*E}\right)$. Note that, for any belief $p < p^{*E}$ both $V_{cf}^E(p)$ and $V_B^E(p)$ have identical learning strategies $\alpha = 0$. Hence, since the confirmatory strategy at p^{*E} has a continuation utility that is dominated by the contradictory learning from source B it follows that $V_B^E(p) \geqslant V_{cf}^E(p)$ for $p < p^{*E}$, which concludes our proof. $\square$

2.A.18 Proof of Lemma 2.7

Proof. The second part of the lemma follows directly from the definition of $\overline{p}^E$. Note that for any belief p, such that $\underline{p}^E < p < \overline{p}^E$, $V_B^E(p)$ ¡ $V_{cf}^E(p)$, as under complete delegation the expert chooses to follow confirmatory learning strategy for these beliefs. In addition, $V_B^E\left(p^{*E}\right) < V_{cf}^E\left(p^{*E}\right)$ and for $p < p^{*E}$, contradictory learning from source B and confirmatory learning have identical strategies. Hence, as updating happens in the same direction and beliefs tend towards p^{*E}, conditional on no signal, it must be the case that $V_B^E(p) < V_{cf}^E(p)$ for all $p < p^{*E}$. As $\underline{p}^E < p^{*E} < \overline{p}^E$. It follows that $V_B^E(p) < V_{cf}^E(p)$ for all $p < \overline{p}^E$. $\square$

2.A.19 Proof of Lemma 2.8

Proof. Strict convexity requires $U_B^{E''}(p) > 0 \; \forall \; p$.

Indeed,

$$U_B^{E''}(p) = \frac{rp^{r-1}(r(\Delta+1))^r\left(\Delta^2 r + \Delta(r(2r+5)+1) - (r+2)\right)\left((p-1)(\Delta(r+1)-(r+2))\right)^{-r-1}}{p-1}$$

Since $r > 0$ and $1 > \Delta > 0$, this expression is positive if and only if

$$\Delta^2 r + \Delta(r(2r+5)+1) - (r+2) < 0 \tag{2.30}$$

Note that for $\Delta = 0$, the above reduces to $-(r+2) < 0$, which clearly holds.

In addition,

$$\frac{d}{d\Delta}[RHS] = 2r\Delta + (1 + r(5+2r)) > 0$$

Hence the RHS of (2.30) increases monotonically with Δ.

Since Δ is bounded above by 1, it suffices to show that (2.30) is satisfied for $\Delta = 1$, i.e. we need to show that

$$r - (r+2) + (r(2r+5)+1) < 0$$

Given our assumptions the above is true if and only if $5 + 4r < \sqrt{33} \rightarrow r < 0.18614$.

Restriction 2.2, requires r to be less than $(2-\Delta)^2/(6\Delta)$.[21] Let $\Delta = 1-\epsilon$, with $\epsilon > 0$. Then, restriction 1 is

$$\frac{(1+\epsilon)^2}{6(1-\epsilon)}$$

Taking the limit as $\epsilon \to 0$

$$\lim_{\epsilon\to 0}\frac{(1+\epsilon)^2}{6(1-\epsilon)} = \frac{1}{6}$$

Given that $1/6 < 0.18614$, (2.30) is satisfied, $U_B^E(p)$ is strictly convex.

To see that $U_B^E(p)$ intersects $U_a(p)$ from above note that by construction $U_B^E(\overline{p}^{*E}) = U_a(\overline{p}^{*E})$. Also, since $U_B^E(p)$ is strictly convex and $U_a(p)$ is linear, it suffices to show that $U_B^{E'}(p)|_{p=\overline{p}^{*E}} < U_a'(p)|_{p=\overline{p}^{*E}}$

Indeed,

$$U_B^{E'}(p)|_{p=\overline{p}^{*E}} - U_a'(p)|_{p=\overline{p}^{*E}} = \frac{(2(r+1)-\Delta)(\Delta+2(r+2))}{(\Delta+1)(\Delta(r+1)-(r+2))}$$

Which is negative since $1 > \Delta > 0$ and $r > 0$.

Finally we need to show that $U_B^E(p) > \max\{U_a(p), U_b(p)\}$ for $p \in [\overline{p}^E, \overline{p}^{*E})$.

By construction $U_B^E(\overline{p}^{*E}) = U_a(\overline{p}^{*E})$ and it is immediate to show that $U_a(\overline{p}^{*E}) > U_b(\overline{p}^{*E})$. As shown above, $U_B^E(p)$ crosses $U_a(p)$ from above, $U_B^E(p)$ is convex and $U_a(p)$ is linear and upward sloping, it follows that $U_B^E(p) > U_a(p)\ \forall p < \overline{p}^{*E}$.

Since, $U_B^E(\overline{p}^{*E}) > U_b(\overline{p}^{*E})$, $U_b(p)$ is linear and downward sloping and $U_B^E(p)$ is strictly convex, $U_B^E(p)$ and $U_b(p)$ can cross at most once if and only if at $p = 0$ $U_b'(p)|_{p=0} - U_B^{E'}(p)|_{p=0} < 0$. This is the case since from $U_B^E(p)$'s strict convexity it will have it's most negative slope at $p = 0$. If at that point the slope is not more negative than $U_b'(p)$ then they cannot cross twice.

Indeed,

[21] See appendix 2.A.13.

$$U_b'(p)|_{p=0} - U_B^{E'}(p)|_{p=0} = -2 - \frac{\Delta+1}{r+1} = -\frac{(1+2r)-\Delta}{r+1} < 0$$

Therefore $U_B^E(p)$ and $U_b(p)$ can cross at most once. It follows that for any p' such that $U_B^E(p') > U_b(p')$, then $U_B^E(p) > U_b(p)\ \forall p \geq p'$.

It remains to show that $U_B^E(\overline{p}^E) > U_b(\overline{p}^E)$. It suffices to show that $\exists\ p' < \overline{p}^E$ s.t. $U_B^E(p') \geq U_b(p')$.

Taking the worst case for the principal, hence $r = r^*$ and $\Delta = 1$ and solving for p;

$$p' = \{p : U_B^E(p) = U_b(p)|_{r=r^*,\ \Delta=1}\} = 0.467785$$

Also, note that $\overline{p}^E|_{r=r^*,\ \Delta=1} = 0.671452$, which concludes the proof.

$\square$

2.A.20 Analytical derivation of $V_i(p), U_i(p), U_{cf}(p)$ and $V_{cf}(p)$ for $i \in \{A, B\}$

To derive analytical expressions for the payoff functions discussed above we use the fact that they must satisfy the HJB derivations for the principal and expert analogously. In addition, we can easily identify the beliefs at which a decision would be taken immediately. Specifically, Lemma 2.2 shows that $\underline{p}^{*P} > \underline{p}^{*E}$ and $\overline{p}^{*P} > \overline{p}^{*E}$. It is immediate that $U_A(p) = U_A^P(p)$, as $\underline{p}^{*P}$ is the belief that P would terminate learning in the first-best scenario. Also, $V_B(p) = V_B^E(p)$ .Therefore, given that equation (2.11) must always be satisfied. $V_A(p)$ is:

$$V_A(p) = \begin{cases} \dfrac{(\Delta+1)p - \frac{(p-1)(\Delta+r(\Delta(\Delta+2r+5)-1)-2)\left(\frac{(\Delta+1)(p-1)r}{\Delta-2r-2}\right)^r\left(\frac{(\Delta+1)pr}{\Delta-2r-2}+p\right)^{-r}}{\Delta+(\Delta-1)r-2}}{r+1} & p > \underline{p}^{*P} \\ V_{b(p)} & p \leq \underline{p}^{*P} \end{cases}$$

$$\overline{p}^{*E} = \frac{r(\Delta+1)}{\Delta-2(r+1)} + 1$$

$$U_B(p) = \frac{\Delta + \frac{(1-p)^{-r}p^{r+1}\left(\frac{(\Delta+1)r}{-\Delta+2r+2}\right)^r(\Delta+r(\Delta(\Delta+2r+5)-1)-2)\left(\frac{\Delta+(\Delta-1)r-2}{\Delta-2r-2}\right)^{-r}}{\Delta+(\Delta-1)r-2} - (\Delta+1)p + 1}{r+1}$$

Under flexible delegation confirmatory learning is defined as the leaning policy were the expert pursues learning as he would under complete delegation. We denote the payoff from this policy by $U_{cf}(p)$ and $V_{cf}(p)$ for the P and E respectively. Hence, $V_{cf}(p) = V_{cf}^{E}(p)$and

$$U_{cf}(p) = \begin{cases} \frac{\Delta-(\Delta+1)p-\frac{p(1-p)^{-r}(\Delta(\Delta+2)r-1)((\Delta+1)p)^{r}}{2r+1}+1}{r+1} & p < p^{*E} \\ \frac{p(\Delta+2)^{r}((\Delta+1)p)^{-r-1}\left((\Delta+1)p(2r+1)\left(\frac{(\Delta+1)p}{\Delta+2}\right)^{r}-(p-1)(\Delta(\Delta+2)(r+1)+1)\left(\frac{1-p}{\Delta+2}\right)^{r}\right)}{(r+1)(2r+1)} & p > p^{*E} \end{cases}$$

Note that by, lemma (2.4), the principal would never find it optimal to intervene at p^{*E}, hence it must be the case that if confirmatory learning is optimal for E at p^{*E}, then it will be part of the equilibrium strategy.

2.A.21 Proof of Lemma 2.9

Proof. The proof proceeds in the following steps. First, note that the α that makes the expert indifferent between learning and taking decision b, henceforth α_{ind}^{E}, can be found using the HJB for the expert and the utility from immediate action b. Namely, we know that it must satisfy both. Then;

$$V(p) = V_b(p) = pv_b^{A} + (1-p)\,v_b^{B},$$

with

$$V'(p) = v_b^{A} - v_b^{B}$$

and satisfies HJB equation

$$rV(p) = p\alpha(p)\left(v_a^{A} - V(p)\right) + (1-p)(1-\alpha(p))\left(v_b^{B} - V(p)\right) - (2\alpha(p)-1)\,p\,(1-p)\,V'(p)\,. \tag{2.31}$$

Given the linearity of $V(p)$, there exits a unique α that satisfies (2.16) and it is given by

$$\alpha_{ind}^{E}(p) = \frac{r-2pr}{\Delta p + 2p}.$$

Observe that $\alpha_{ind}^{E}(p) < \alpha^{*}(p)$,

$$\alpha_{ind}^{E}(p) = \frac{r-2pr}{\Delta p + 2p} < \frac{r(1+\Delta-2p)}{p(2-\Delta)} = \alpha^{*}(p)$$

$$\Delta\left(\Delta^{2}-4\right)pr(-\Delta+4p-4) > 0$$

Which always holds as $(\Delta^2 - 4)$ and $(-\Delta + 4p - 4)$ are negative.

The linearity of the HJB in α implies that when the optimal α in the experts decision problem is 1 then his utility is increasing in α. Therefore, any $\alpha' > \alpha^E_{ind}(p)$ implies $V(\alpha') > V_{ind}$. □

2.A.22 Proof of Corollary 2.1

Proof. By Lemma 2.4, $r \leq r^* = \frac{2-\Delta}{2\Delta(\Delta+3)}$ As

$$\frac{dr^*}{d\Delta} = \frac{\Delta^2 - 4\Delta - 6}{2\Delta^2(\Delta+3)^2} < 0, \quad \text{for } \Delta \in (0,1)$$

The highest value of r^* is

$$\lim_{\Delta \to 0^+} r^* = \frac{1}{8}$$

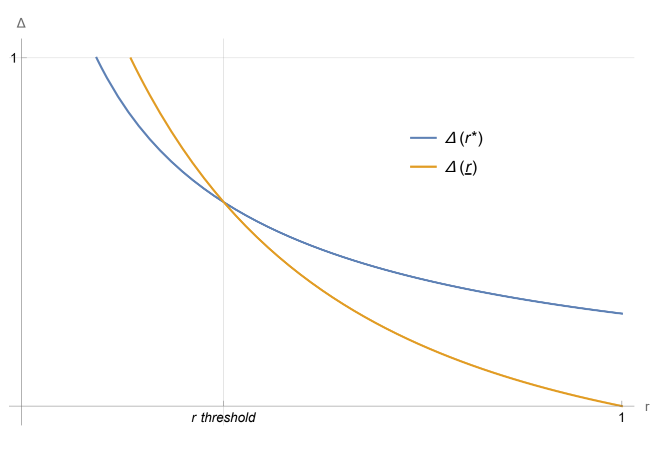
Further, note that by (2.10), $\underline{r}(\Delta)$ must satisfy the following condition.

$$\Delta = \frac{-r(2r+1)^{\frac{1}{r}} + r + 2}{r+1}$$

Also, by Lemma 2.3, $\underline{r}(\Delta)$ decreases in Δ, therefore setting $\Delta = 1$ above, yields The inverse of $r^*(\Delta)$ is:

$$\Delta(r^*) = \frac{\left(\sqrt{\frac{1}{r^2} + \frac{28}{r} + 36} - 6\right) r - 1}{4r} \tag{2.32}$$

Comparing, equation (2.10) and the inverse of $r^*(\Delta)$.



As can be seen from the graph above, there exists a threshold for r below which both confirmatory and contradictory strategies must be part of the equilibrium as the highest Δ dictated by r^* is lower than the lowest Δ required for contradictory strategy to be optimal only.

$\square$

2.A.23 Verification of flexible delegation equilibrium

No learning

If $\bar{r}(\Delta) < r(\Delta)$ then the principal's boundary condition 2.8 binds and $y = \text{argmax } U_x(p)$.

Contradictory learning only

If $\bar{r}(\Delta) > r(\Delta) > \underline{r}(\Delta)$ then by Lemma 2.1 both the expert's and the principal's optimal learning strategy if they both had access to the learning technology would be contradictory only.

Case 1 : $V_A\left(\underline{p}^{*P}\right) \geq V_B\left(\underline{p}^{*P}\right)$

If $V_A\left(\underline{p}^{*P}\right) \geq V_B\left(\underline{p}^{*P}\right)$ the expert learns from source A for beliefs $p < \check{p}^E$ and the principal will take decision b at belief $\underline{p}^{*P}$ conditional on no signal arriving until $\underline{p}^{*P}$ is reached. For $p > \check{p}^E$ the expert will learn from source B . If no signal arrives in this case, the expert will take decision a at belief $\overline{p}^{*E}$.

$$Y^* = \begin{cases} b & p \leq \underline{p}^{*P} \\ \emptyset & p > \underline{p}^{*P} \end{cases}$$

$$(\alpha^*, X^*) = \begin{cases} (\emptyset, b) & p \leq \underline{p}^{*E} \\ (1, \emptyset) & \underline{p}^{*E} < p \leq \check{p}^E \\ (0, \emptyset) & \check{p}^E < p < \overline{p}^{*E} \\ (\emptyset, a) & p \geq \overline{p}^{*E} \end{cases}$$

Proof. To show that this constitutes an equilibrium the conditions below must be satisfied.

$$\max\left\{\max_{\alpha} H^E(\alpha) - rV(p),\ V^E(p) - V(p)\right\} = 0 \tag{2.33}$$

$$\max\left\{H^P(\alpha^*) - rU(p),\ U^P(p) - U(p)\right\} = 0 \tag{2.34}$$

with

$$H^E(\alpha) \equiv p\alpha(v_a^A - V(p)) + (1-p)(1-\alpha)(v_b^B - V(p)) - (2\alpha - 1)p(1-p)V'(p)$$

and

$$H^P(\alpha) \equiv p\alpha^*(u_a^A - U(p)) + (1-p)(1-\alpha^*)(u_b^B - U(p)) - (2\alpha^* - 1)p(1-p)U'(p)$$

For the expert note that for $p \in [0, \underline{p}^{*E}] \cup (\underline{p}^{*P}, 1]$ the expert pursues his first best strategy. It therefore suffices to show that condition (2.33) is met at $p \in (\underline{p}^{*E}, \underline{p}^{*P}]$. By Lemma 2.2 this region is non-empty.

Clearly, $V^E(p) = V(p)\ \forall p \in (\underline{p}^{*E}, \underline{p}^{*P}]$, as given the principal's optimal strategy, the expert derives the utility from immediate learning in this region. Moreover, any policy α does not affect the experts utility given the principal's strategy as no learning takes place. Hence (2.33) is satisfied in all points of differentiability.

From the perspective of the principal, we can split this into two regions $p \leq \check{p}^E$ and $p > \check{p}^E$. For $p < \check{p}^E$ the principal pursues and obtains her first best. Therefore it remains to show that $U_B^E(p) > \max\{U_a(p),\ U_b(p)\}$ for $p \in (\check{p}^E, 1]$.

By construction, $U_0^E(\overline{p}^{*E}) = U_a(\overline{p}^{*E}) > U_b(\overline{p}^{*E})$. Also, by Lemma 2.8 $U_B^E(p)$ intersects $U_a(p)$ from above. From Lemma 2.8 we know that $U_B^E(p)$ is convex and since $U_a(p)$ is linear, $U_B^E(p) > U_a(p),\ \forall p < \overline{p}^{*E}$.

□

Case 2: $V_A\left(\underline{p}^{*P}\right) < V_B\left(\underline{p}^{*P}\right)$

The expert will not pursue contradictory learning from source A, as doing so is dominated by pursuing contradictory learning from source B. By Lemma 2.5 $\check{p}^E \leqslant p^{*E}$, and by Lemma 2.6, the expert will pursue contradictory learning from source B only. However, at belief $\widehat{p}$, with $\widehat{p} = \{p|U_b(p) = U_B(p)\}$, the expert will increase α, so as to make the principal indifferent between taking $y = \emptyset$ and $y = b$. The expert will keep the principal indifferent up to $\underline{p}^{*P}$, at which point the boundary condition for the expert will hold and decision b will be taken. If no signal arrives, decision a will be taken by the expert at $\overline{p}^{*E}$.

$$Y^* = \begin{cases} b & p \leq \underline{p}^{*P} \\ \emptyset & p > \underline{p}^{*P} \end{cases}$$

$$(\alpha^*, X^*) = \begin{cases} (\emptyset, b) & p \leq \underline{p}^{*E} \\ (1, \emptyset) & \underline{p}^{*E} < p \leq \underline{p}^{*P} \\ (\alpha^*(p), \emptyset) & \underline{p}^{*P} < p \leq \widehat{p} \\ (1, \emptyset) & \widehat{p} < p \leq \check{p}^{E} \\ (0, \emptyset) & \check{p}^{E} < p < \overline{p}^{*E} \\ (\emptyset, a) & p \geq \overline{p}^{*E} \end{cases}$$

Proof. We need to show that conditions (2.34) and (2.33) are satisfied at all points of differentiability.

(2.34) is satisfied for $p < \hat{p}$ by construction as $U(p) = U_b(p) > U_a(p)\ \forall p \leq \hat{p}$. For $p \in (\hat{p}, \overline{p}^{*E})$, we need to show that $U_0^E(p) > \max\{U_b(p), U_a(p)\}$. The argument supporting this is identical to the one given in case 1 above.

(2.33) is satisfied for $p \geq \hat{p}$ as in this region the expert pursues and attains his first best. Further, an identical argument to that made in Case 1 above shows that (2.33) is also satisfied for $p \leq \underline{p}^{*P}$. Therefore it remains to show that it is also satisfied for $p \in (\underline{p}^{*P}, \hat{p})$. This must be the case by construction. We know that in this region the expert would like to pursue a learning strategy with $\alpha = 1$. Given that the experts HJB is linear in α, this means that in this region the HJB is increasing in α. Therefore the maximum the expert can attain conditional on not being stopped by the principal is α^* by it's definition. Lemma 2.9 ensures that the expert is better off learning with intensity α^* than taking immediate action.

□

Confirmatory and Contradictory learning

If $r(\Delta) < \underline{r}(\Delta)$ then both the expert's and the principal's optimal learning strategies would include both confirmatory and contradictory learning by Lemma 2.1.

Case 1: $V_A(\underline{p}^{*P}) \geq V_{cf}(\underline{p}^{*P})$

In this case then the principal will take decision b at any p below $\underline{p}^{*P}$. For $p > \underline{p}^{*P}$ the expert will learn from source A until $p = p''$, with $p'' = \left\{p | V_A^P(p) = V_{cf}^E(p)\right\}$. For beliefs $\underline{p}'' < p < \overline{p}^E$ the expert will pursue confirmatory learning strategy and for $p > \overline{p}^E$ the expert will learn from source B until $\overline{p}^{*E}$, where decision a will be taken by the expert.

$$Y^* = \begin{cases} b & p \leq \underline{p}^{*P} \\ \emptyset & p > \underline{p}^{*P} \end{cases}$$

$$(\alpha^*, X^*) = \begin{cases} (\emptyset, b) & p \leq \underline{p}^{*E} \\ (1, \emptyset) & \underline{p}^{*E} < p \leq \underline{p}^E \\ (0, \emptyset) & \underline{p}^E < p < p^{*E} \\ \left(\frac{1}{2}, \emptyset\right) & p = p^{*E} \\ (1, \emptyset) & p^{*E} < p \leq \overline{p}^E \\ (0, \emptyset) & \overline{p}^E < p < \overline{p}^{*E} \\ (\emptyset, a) & p \geq \overline{p}^{*E} \end{cases}$$

Proof. We need to show that conditions (2.34) and (2.33) are satisfied at all points of differentiability.

The expert pursues and attains his first best for $p \in (\underline{p}^E, 1]$, hence at these beliefs (2.33) must be satisfied. For $p \in [0, \underline{p}^E]$ the arguments showing that (2.33) is satisfied are identical to case 1 when contradictory learning only is pursued.

The principal for $p \in [0, \underline{p}^E]$ pursues and attains her first best, hence (2.34) is satisfied.

Further, for $p \in [\bar{p}^E, \bar{p}^{*E}]$ we need to show that $U_0^E(p) > \max\{U_a(p), U_b(p)\}$. By Lemma 2.8, $U_B^E(p)$ intersects $U_a(p)$ from above. From Lemma 2.8 we know that $U_B^E(p)$ is convex and since $U_a(p)$ is linear, $U_B^E(p) > U_a(p) > U_b(p),\ p \in [\bar{p}^E, \bar{p}^{*E}$.

It therefore remains to show that (2.34) is satisfied for $p \in (\underline{p}^E, \bar{p}^E)$. By Lemma 2.4 $U_{cf}^E(p) \geq \max\{U_a(p), U_b(p)\}$ for $p \in (\underline{p}^E, \bar{p}^E)$ (2.34) is satisfied. ☐

Case 2: $V_A\left(\underline{p}^{*P}\right) < V_{cf}\left(\underline{p}^{*P}\right)$

If $V_A\left(\underline{p}^{*P}\right) < V_{cf}\left(\underline{p}^{*P}\right)$, then the expert will not pursue contradictory learning from source A. The principal will take decision b for any $p \leq \underline{p}^{*P}$. The expert will keep the principal indifferent for $\underline{p}^{*P} < p < \widetilde{p}$ and by Lemma 2.7, the expert will find it optimal to pursue confirmatory learning for beliefs $\widetilde{p} \leq p < \overline{p}^{E}$, with $\widetilde{p} = \left\{p|U^{E}_{cf}(p) = U_b(p)\right\}$. Moreover, the expert will pursue contradictory learning from source B for beliefs $p > \overline{p}^{E}$. At $\overline{p}^{*E}$decision a will be taken by the expert.

$$Y^* = \begin{cases} b & p \leq \underline{p}^{*P} \\ \emptyset & p > \underline{p}^{*P} \end{cases}$$

$$(\alpha^*, X^*) = \begin{cases} (\emptyset, b) & p \leq \underline{p}^{*E} \\ (1, \emptyset) & \underline{p}^{*E} < p \leq \underline{p}^{*P} \\ (\alpha^*(p), \emptyset) & \underline{p}^{*P} < p \leq \widetilde{p} \\ (0, \emptyset) & \widetilde{p} < p < p^{*E} \\ \left(\frac{1}{2}, \emptyset\right) & p = p^{*E} \\ (1, \emptyset) & p^{*E} < p \leq \overline{p}^{E} \\ (0, \emptyset) & \overline{p}^{E} < p < \overline{p}^{*E} \\ (\emptyset, a) & p \geq \overline{p}^{*E} \end{cases}$$

Proof. We need to show that conditions (2.34) and (2.33) are satisfied at all points of differentiability.

For the principal for $p \in (\widetilde{p}, 1]$ by Lemma 2.8 and Lemma 2.4 this is the case. For $p \in [0, \widetilde{p}]$, (2.34) is satisfied by construction since $U(p) = U_b(p) > U_a(p)$.

For the expert, for $p \in (\widetilde{p}, 1]$ the expert pursues and attains his first best hence (2.33) is satisfied. For $p \leq \widetilde{p}$ This must be the case by construction. We know that in this region the expert would like to pursue a learning strategy with $\alpha = 0$. Given that the experts HJB is linear in α, this means that in this region the HJB is decreasing in α. Therefore the maximum the expert can attain conditional on not being stopped by the principal is α^* by it's definition. Lemma 2.9 ensures that the expert is better off learning with intensity α^* than taking immediate action. □

Possible Indifferences

Contradictory learning only

Case 1 : $V_A\left(\underline{p}^{*P}\right) \geq V_B\left(\underline{p}^{*P}\right)$

Possible indifference from $p' = \{p|V_B^E(p) = V_A^P(p)\}$ until $p'' = \{p|U_b(p) = U_B^E(p)\}$.

Case 2: $V_A\left(\underline{p}^{*P}\right) < V_B\left(\underline{p}^{*P}\right)$

Possible indifference from $\underline{p}^{*P}$ until $p' = \{p|U_b(p) = U_B^E(p)\}$.

Confirmatory and Contradictory learning

Case 1: $V_A\left(\underline{p}^{*P}\right) \geq V_{cf}\left(\underline{p}^{*P}\right)$

Possible indifference from $p' = \{p|V_A^P(p) = V_{cf}^E(p)\}$ until $p'' = \{p|U_{cf}(p) = U_b(p)\}$.

Case 2: $V_A\left(\underline{p}^{*P}\right) < V_{cf}\left(\underline{p}^{*P}\right)$

Possible indifference from $\underline{p}^{*P}$ until $p' = \{p|U_b(p) = U_{cf}(p)\}$.

2.A.24 Proof of Theorem 2.1

Proof. We start by proving the second part of the theorem. If $V_A(\underline{p}^{*P}) > V_B(\underline{p}^{*P})$ for $r(\Delta) \in \left(\underline{r}(\Delta),\ \bar{r}(\Delta)\right)$ or $V_A(\underline{p}^{*P}) > V_{cf}(\underline{p}^{*P})$ for $r(\Delta) \in \left(0,\ \underline{r}(\Delta)\right)$, then flexible delegation equilibrium shows that the principal never intervenes for any $p > \underline{p}^{*P}$ and the expert pursues the same learning policy as he does under complete delegation. Therefore, the utility derived by the principal between flexible and complete delegation is identical for all $p > \underline{p}^{*P}$. However, for $p \in [\underline{p}^{*E}, \underline{p}^{*P}]$ under complete delegation the expert pursues learning, while under flexible delegation decision b is taken by the principle. By the definition of $\underline{p}^{*P}$, the principal's utility is higher by no learning as opposed to any learning in that region, which makes her strictly worse of under complete delegation.

It remains to show that there exists a $\Delta(r)$ above which $V_A(\underline{p}^{*P}) > V_B(\underline{p}^{*P})$ and $V_A(\underline{p}^{*P}) > V_{cf}(\underline{p}^{*P})$.

Consider the case that $\Delta(r) = 0$, in this case the preferences of the expert and the principal are fully aligned and $\underline{p}^{*P} = \underline{p}^{*E}$. Che and Mierendorff (2019) show that

$V_A(\underline{p}^{*E}) < V_B(\underline{p}^{*E})$ for $r(\Delta) \in \Big(\underline{r}(\Delta), \bar{r}(\Delta)\Big)$ and $V_A(\underline{p}^{*E}) < V_{cf}(\underline{p}^{*E})$ for $r(\Delta) \in \Big(0, \underline{r}(\Delta)\Big)$.

Further, as $\Delta(r)$ increases both $V_B(p)$ and $V_{cf}(p)$ continue to be identical to the expert's first-best, but $V_A(p)$ monotonically decreases for all p, since $\underline{p}^{*P} - \underline{p}^{*E}$ increases with $\Delta(r)$. Therefore, it must the case that for some $\Delta(r) \equiv \underline{\Delta}(r)$ and $r(\Delta) \in \Big(\underline{r}(\Delta), \bar{r}(\Delta)\Big)$, $V_A(\underline{p}^{*P}) = V_B(\underline{p}^{*P})$ and for some $\Delta(r) \equiv \underline{\Delta}(r)$ and $r(\Delta) \in \Big(0, \underline{r}(\Delta)\Big)$, $V_A(\underline{p}^{*P}) = V_{cf}(\underline{p}^{*P})$. For $\Delta(r) > \underline{\Delta}(r)$, the second part of the theorem holds.

To show that the first part of the theorem holds, note that for $\Delta(r) < \underline{\Delta}(r)$, the flexible delegation equilibrium learning policy suggests that for some belief $p \in \Big(\underline{p}^{*P},\ \underline{p}^{*P} + \epsilon\Big]$ the expert will pursue a policy which makes the principal indifferent between making a decision and taking immediate action b. Further, note that $\Delta(r) = 0$ implies that $\underline{p}^{*P} = \underline{p}^{*E}$ as above.

Let $r(\Delta) \in \Big(\underline{r}(\Delta),\ \bar{r}(\Delta)\Big)$ and $\Delta(r) < \underline{\Delta}(r)$.

Then,

$$U_B(\underline{p}^{*E}) = U_B(\underline{p}^{*P}) > U_b(\underline{p}^{*P})$$

Under flexible delegation, there exists ϵ such that $U_{fl}(\underline{p}^{*P} + \epsilon) = U_b(\underline{p}^{*P} + \epsilon)$, where the subscript fl denotes the equilibrium payoff under flexible delegation, where the expert is pursuing the learning that makes the principal indifferent. Under complete delegation, for $\epsilon \to 0$ the expert would pursue contradictory with $a = 1$ and the principal would derive utility $U_B(\underline{p}^{*P} + \epsilon)$.

Therefore, taking the limit as $\epsilon \to 0$ and $\Delta(r) \to 0$. It must the the case that

$$U_B(\underline{p}^{*P} + \epsilon) > U_b(\underline{p}^{*P} + \epsilon)$$

Where the right side corresponds to the utility derived by the principle under complete delegation and the left side that derived by the principal under flexible delegation.

A very similar argument can be made for the case that $r(\Delta) \in \Big(0,\ \underline{r}(\Delta)\Big)$.

$\square$

2.A.25 Heuristic derivation of the HJB equation

By the dynamic programming principle (DPP), in discrete time, with a period length $h > 0$, the expert's value function $V(p)$ would be characterized by the Bellman equation:

$$V(p) = \max_{\alpha} \left\{ \underbrace{p\left(1-e^{-\alpha h}\right) v_a^A + (1-p)\left(1-e^{-(1-\alpha)h}\right) v_b^B}_{\text{payoff if signal arrives during the period}} + e^{-rh} \underbrace{\left(pe^{-\alpha h} + (1-p) e^{-(1-\alpha)h}\right)}_{\text{probability no signal arrives}} \underbrace{V\left(p^{\alpha h}\right)}_{\text{continuation payoff}} \right\},$$

where α is the allocation of the learning intensity that remains fixed for the duration of the period and $p^{\alpha h}$ is the revised belief at the end of the period. By Taylor approximation, when h is small,

$$1-e^{-\alpha h} \approx \alpha h \text{ and } 1-e^{-(1-\alpha)h} \approx (1-\alpha) h$$
$$e^{-rh} \approx 1 - rh.$$

So

$$V\Big(p\Big) \geq \max_{\alpha} \Bigg\{ \Bigg(p\alpha v_a^A + \Big(1-p\Big)\Big(1-\alpha\Big) v_b^B \Bigg) h + \Big(1-rh\Big)\Bigg(1-\Big(p\alpha + \Big(1-p\Big)\Big(1-\alpha\Big)\Big)h\Bigg) V\Big(p^{\alpha h}\Big) \Bigg\}. \tag{2.35}$$

If V is differentiable, $V\left(p^{\alpha h}\right)$ can be approximated as

$$V\left(p^{\alpha h}\right) \approx V(p) - (2\alpha - 1)\, p\, (1-p)\, V'(p)\, h$$

Omitting the terms of order h and smaller in (2.35) results in

$$0 \geq \max_{\alpha} \left\{ p\alpha \left(v_a^A - V(p)\right) + (1-p)(1-\alpha)\left(v_b^B - V(p)\right) - (2\alpha - 1)\, p\, (1-p)\, V'(p) - rV(p)] \right\}.$$

Chapter 3

Endogenous Systemic Risk

3.1 Introduction

Systemic risk taking behaviour of banks, thought of as the risk taking behaviour of banks that incorporates the systemic consequences of their individual decisions is generally considered an externality in most of the economic literature. In other words, although it is now evident that individual risk taking of banks may have systemic consequences, this is not something that any one individual bank is usually considered to internalize in their decision making process. There are many reasons why such a view may be held and also be correct. For example, in very complex and dense networks of financial actors, it is often the case that the behaviour of one individual bank does not have any systemic consequence.

However, during the financial crisis of 2007/08, it became evident that regulators and central banks, needed to make drastic decisions on the basis of the systematic consequence of individual bank failures. Arguments such as too big to fail, or too interconnected to fail, have been extensively used as a description of why it was necessary to provide liquidity to insolvent institutions, which although should have been allowed to go bankrupt if viewed individually, had to be bailed out when viewed as part of a bigger network.

This being the case, it seems natural to expect that under certain conditions, it is possible for individual banks to internalize the decision making process of the regulators, and internalize the systemic risk that results from individual decision making. In this paper, we construct a simple model of banks, connected through sharing investment opportunities. We characterise the conditions under which there exists a unique Nash Equilibrium in which banks endogenise the systemic implications of their actions thereby increasing systemic risk. Our model is different from the existing literature as we study the possibility of banks increasing their systemic importance by sharing portfolios, rather than by taking correlated risks, which has often been assumed in the literature. Although very stylized, our model suggests that risk-neutral, profit-maximizing banks may be willing to share investments that increase their systemic importance, even at the detriment of expected profits. In our model this will only happen when banks ensure that the regulator will always intervene if required. Crucially, this result depends on the regulator's policy rule, which in our case is the extreme rule that the regulator will never intervene unless not doing so will result in systemic collapse.

The next section discusses the literature on systemic risk and bailouts, section 3.3 develops the baseline model, section 3.4 augments the baseline model by adding depositors and discusses our main results. Section 3.5 concludes.

3.2 Literature Review

Prior to the 2007/08 financial crisis the literature mostly focused on systemic risk as an externality. Freixas, Parigi, and Rochet (2000) model systemic risk in the interbank market in a model with depositors. They find that the interbank market helps reduce the liquidity holdings of banks and increases the resilience of the banking system as liquidity shocks can be absorbed. However, it is also susceptible to speculative gridlocks, stemming from depositors' expectations about the liquidity of banks. In their set up if the central bank is faced with an insolvent bank, its optimal response is to allow for that bank to close, while providing the rest of the banks the liquidity they will lose due to the closure. Their model, which is a simplified version of Diamond and Dybvig (1983), with only risk neutral, patient depositors, is based on the idea that depositors 'travel' and can demand to withdraw from any bank.

More recently, internalizing the systemic consequences of risk taking has become more commonplace. For example, Dell and Ratnovski (2012) are interested in the relationship between risk taking of banks and bailouts. In their model, the expectation of a bailout also encourages increased risk taking. However, in their two-bank set-up they model contagion as the probability of bank j failing conditional on bank i failing, independently from each bank's exposure to risk. They find that bailouts increase moral hazard but when the risk of contagion is high the prospect of a bailout decreases the risk taking behaviour of banks. Differently from our setting, there is no interbank lending in their model.

In a more complicated set up, Farhi and Tirole (2012), develop an argument that is close in spirit to ours. In their paper, banks' private leverage decisions depend on the expected policy response if they require some form of liquidity. In their model, there is an incentive for banks to need liquidity when the whole financial system is in distress, rather than on our own, as the former will motivate the regulators to implement beneficial policies. This logic is very close to ours, however, we are interested in how banks provide liquidity to each other, rather than in the external risks they are exposed to. In a similar vein, Acharya (2009) designs bank regulation when banks can choose the correlation of returns on their assets. The contribution of their paper is to describe optimal regulator policies when risk correlation is an endogenous variable. They suggest that bailouts accompanied by dilution of owner's equity is ex-ante optimal when banks fail, particularly if both banks fail together. In a similar set up, with two banks in two regions, Dasgupta (2004) models contagion stemming from banks insuring against regional depositors' liquidity shocks using the interbank market. Due to the risk of contagion they show that banks never fully insure in equilibrium.

Erol (2017) study endogenous network formation when firms anticipate that their level of connectedness can induce bailouts when distressed. They show that the expectation of bailouts, makes firms less concerned about the decisions of the firms they are connected with, which increases connectedness and systemic risk. This result has some similarities with our intuition that banks would be willing to share less profitable investments if this increased their systemic importance. In addition, the existence of intervention creates 'network hazard', i.e. core-periphery structures in which the core firms are systemically important and can amplify contagion when insolvent.

The paper that is closest to ours is Acharya and Yorulmazer (2007). Like us they observe that crisis prevention conflicts with crisis management. The central bank is more likely to intervene when many banks fail, therefore, creating an incentive for risk correlation instead of risk diversification from banks. If few banks fail, surviving banks are able to step in and buy failing ones, thus not requiring the CB to intervene. Banks herd ex-ante, for example by investing in similar industries or similar financial assets, in order to increase the likelihood of intervention, which leads to systemic banking crises. In contrast to the too big-to-fail argument, this highlights the systemic role of small banks as well as large ones. Their analysis indicates that this problem is more severe when, a) agency problems are high (eg. fraud by bank owners) and b) in times where fiscal costs of bailing out banks are high. When modelling banks of different sizes, big banks have an incentive to diversify risk. However, they argue that the fact that when many banks fail at the same time regulators have to respond, creates incentives for small banks to herd by investing in similar industries. Further, Acharya, Mehran, and Thakor (2016) investigate optimal banks' capital structures and show that in the presence of a regulator who cares about systemic risk there exists an equilibrium in which banks are over-leveraged in order to fund correlated inefficiently risky loans. This creates a double moral hazard since banks invest in bad projects and have too much equity as opposed to debt, which leads to insufficient loan monitoring. In their framework requiring banks to maintain a 'special capital account', unavailable to creditors, reduces these effects. Crucially, in our paper we consider that banks connect via the inter-banking market rather than by correlated investments. This enables us to analyse the network effects of the banks' incentives.

Cabrales, Gottardi, and Vega-Redondo (2017) study the properties of financial networks in which connections help firms share risk but increase the possibility of contagion. Similarly to our setting in their model financial firms have access to their own risky project but can also invest to those of the firms they are connected to. They characterize network structures that maximize social surplus, defined as the structure that minimizes the expected number of defaults. They find that socially optimal networks depend on the type of risk firms are exposed to. Put coarsely, if the likelihood of a big shock is high then network structures that protect from contagion are optimal - these are structures that exhibit 'maximal segmentation'. Conversely, highly connected networks that allow for maximal risk sharing are optimal. In our paper we analyse a very similar risk sharing and contagion transmission mechanism but we incorporate the idea that firms also endogenise the regulators response also.

3.3 Model

3.3.1 Banks

There exist N banks each of which has exclusive access to an individual investment opportunity. This opportunity can be thought of as being an external investment, or a portfolio of investments, that is unique to each individual bank. For simplicity, each opportunity requires at maximum one unit of capital. Banks are risk neutral and profit maximizers. This simply means wanting to achieve the highest expected return from their investment. For the benchmark model, each bank is endowed with one unit of capital, this assumption will be relaxed in section 3.4 when depositors will be introduced into the framework but is currently maintained in order to demonstrate the equilibrium behaviour of banks with greater clarity.

ShickIn period 0, banks choose to invest their capital on their own project or can choose to invest any part of their capital in projects of other banks, provided that there is mutual agreement. In order to simplify the analysis, returns on individual projects are either high or low. In particular, each bank can be of type $\theta \in \{l, h\}$, depending on the return of the portfolio they own. The proportion of high type is common knowledge and denoted by μ. Return of bank i is $R_i \in \{R^\theta, 0\}$, with $R^h > R^l$. Further, $pr(R^\theta) = p \; \forall \; i$. Ω is the state of the world, i.e. the resolution of all uncertainty, with $\omega = \{R_1, \ldots, R_N\}$. We also assume that $pR^l > 1$, so that for each bank investing in its own portfolio is ex-ante optimal. Banks know their type but not that of other banks. Moreover, they have the option of not investing all of their collected capital and keep as much as they want as reserve, in case their project is distressed. In this case, the reserved capital, yields 0 excess return. In the benchmark model, since $pR^l > 1$, it will never be individually rational for banks to maintain any reserve capital, however, this may not be the case when depositors are added to the model.

In addition, with probability $1 - p$ the investment opportunity of a bank is distressed in period 1. The shock is i.i.d to each project and the likelihood is known. In this case, the bank will still make a positive return from the investments it has on other projects, assuming these projects are not distressed also. In the benchmark model, an assumption that will be relaxed in the general case, we consider a bank to be distressed and require liquidity if the shock destroys more than a threshold level β, henceforth bankruptcy threshold, of the initial capital of the bank. In this case, the bank is considered bankrupt, unless liquidity is provided by the CB. Following liquidity provision, the returns to the bank remain intact.

Figure 3.3.1: Timeline of benchmark model

$t = 0$	$t = 1$	$t = 2$
• Banks choose how to share investment projects	• Ω is fully revealed • CB decides whether to provide liquidity to banks	• Returns are realized

Effectively, this shock represents external factors to the investment that may affect its liquidity. This may also be used to capture certain dynamics during a recession, the shock could represent something like the effect of an overall lack of liquidity within the economy. As a consequence, a solvent project may be affected by illiquid partners and therefore require liquidity from the bank in order to survive.

Under these assumptions, each bank i maximizes its profits Π_i;

$$\max_{\mathbf{s}_{ij}} \mathbb{E}[\Pi_i] = \mathbb{E}\left[\sum_{j=1}^{N} s_{ij} R_j | L\right] = p \sum_{j=1}^{N} s_{ij} R_j + (1-p) \sum_{j=1}^{N} \mathbb{1}_{L_j=1} s_{ij} R_j \tag{3.1}$$

subject to:

$$\sum_{j=1}^{N} s_{ij} \leq 1,$$

where s_{ij} is the equity share of bank i in bank j's portfolio. Any sharing s_{ij} has to be agreed by both bank i and bank j. We denote Π_i^1 the period 1 returns on bank i, before the CB has decided whether to provide liquidity. The set of all period 1 returns for all banks is denoted by $\mathbf{\Pi}^1$. In addition, we define;

$$L_i : \mathbf{\Pi}^1 \rightarrow \{0, 1\}$$

the function taking the value of 1 if the CB decides to provide liquidity to bank i and 0 otherwise.

3.3.2 Central Bank

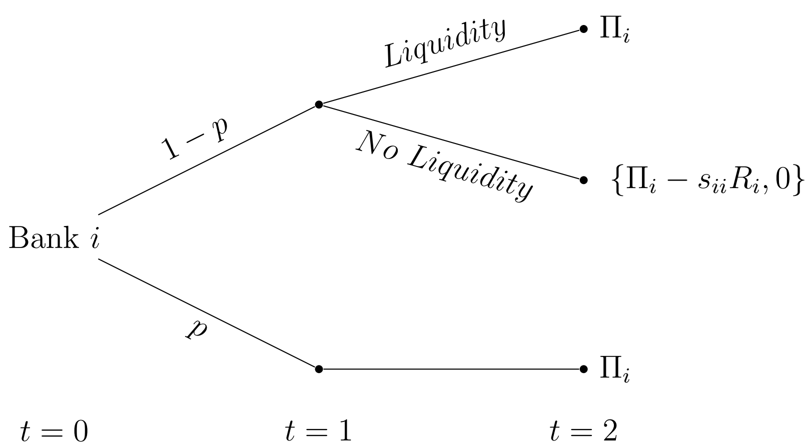
To complete the benchmark model, we need to specify the role of the CB and characterize the decision rule for providing liquidity to distressed banks.

In period 1, the central bank decides whether to save the distressed banks or not. The CB has by assumption unlimited funds. In addition, in period 1 all uncertainty has been resolved, and all aspects of the environment are common knowledge. The CB operates under a simple rule: Provide liquidity to any distressed bank, if and only if not doing so will cause a collapse of the banking system. Hence, in our framework, systemic failure is defined as the extreme scenario in which all banks require liquidity in order to survive. The CB will provide liquidity to a distressed bank i when not doing so creates a systemic collapse.

Definition 3.1 (CB Rule)**.** *The CB will provide liquidity to bank i if and only if not doing so implies that*

$$\sum_{j=1,j\neq i}^{N} \Pi_j = 0$$

Figure 3.3.2: Timeline of benchmark model for bank i - No sharing



3.3.3 Equilibrium Characterization

To begin the equilibrium characterization, consider the case in which there is no central bank intervention. Then our model reduces to a static game in which banks need to choose how much to share with each other.

Lemma 3.1. *In the absence of CB intervention, there is no NE in which high type banks share their portfolios.*

Proof. If no CB, then bank i maximizes;

$$\max_{\mathbf{s}_{ij}} p \sum_{j=1}^{N} s_{ij}\mathbb{E}[R_j] \tag{3.2}$$

subject to:

$$\sum_{j=1}^{N} s_{ij} \leq 1,$$

Further, if banks do not know the other's type, for any high type bank i,

$$\mathbb{E}[R_j] = \frac{(\mu N - 1)R^h + (1-\mu)NR^l}{N-1} < R^h.$$

and for any low type bank i,

$$\mathbb{E}[R_j] = \frac{(\mu N)R^h + ((1-\mu)N - 1)R^l}{N-1} > R^l.$$

Therefore, the low type banks always have an incentive to mimic a high type. As there is no way to credibly convey one's type, high type banks prefer not to share their investment.

□

Corollary 3.1. *In the absence of a CB, in any NE any low type bank i is indifferent between sharing its investment or investing in its own portfolio only.*

Proof. By lemma 3.1 in any NE high type banks invest in their own portfolios only. Therefore, $\mathbb{E}[R_j] = R^l$. As banks are risk neutral and the expected returns are the same for all banks willing to share, they are indifferent between any s_{ij} that is accepted by the other bank. □

Lemma 3.2. *In any pooling equilibrium, i.e. an equilibrium in which high and low type banks behave indistinguishably,*

$$s_{ij} \leq \beta, \ \forall i \neq j.$$

Proof. From lemma 3.1 we know that in the absence of CB intervention this is satisfied. Consider the case that there is CB intervention and assume that there exists an equilibrium in which $s_{ij} > \beta$. Since there is CB intervention, bank j failing implies that

$$\sum_{i=1, i \neq j}^{N} \Pi_i = 0$$

by definition. However, bank i can propose a different level of sharing $\beta \leq s'_{ij} < s_{ij}$. This level of sharing, would not impact any other bank, and therefore would not affect the CB decision to provide liquidity to j. In addition, it increases the expected payoff of both i and j if they are high types so it will be accepted by all high type banks. Therefore, if a bank rejects this offer, then this must be a low type bank and this cannot be a pooling equilibrium.

□

Lemma 3.3. *For banks to be fully protected then each bank i has to invest in a single other bank j, with*

$$s_{ij} = \beta.$$

Proof. Assume not. Hence assume that $s_{ij} < \beta$. Then, if only bank j's portfolio collapses, i.e. the rest of the banks' portfolios have positive returns, then bank i's shock is less than β. Hence, the CB will not intervene. □

Full protection implies that the banking system is robust to even a single distressed portfolio. Having established how banks can be fully protected, the remaining results show that if banks choose to be protected they will do so only if they are fully protected.

Lemma 3.4. *For any bank i with a high investment opportunity that chooses to share its investment,*

$$\sum_{j=1, j \neq i}^{N} s_{ij} = \beta.$$

Proof. If $\sum_{j=1, j \neq i}^{N} s_{ij} < \beta$ then the CB will only provide liquidity to bank i if its own project fails. Therefore bank i's chance of CB intervention is not helped by its investments to the other banks. As the returns of its own portfolio is higher than the expected return of its investment in other banks' portfolios, bank i would be better off investing only in itself. Similarly, if $\sum_{j=1, j \neq i}^{N} s_{ij} > \beta$ bank i can reduce its holdings until $\sum_{j=1, j \neq i}^{N} s_{ij} = \beta$ and maintain the same probability of liquidity provision while increasing its expected profits. □

Corollary 3.2. *Any bank i with a high investment opportunity that chooses to share its investment, will do so if and only if it results in full protection.*

Proof. This is a direct result from Lemmas 3.3 and 3.4.

By Lemma 3.4

$$\sum_{j=1, j\neq i}^{N} s_{ij} = \beta.$$

By Lemma 3.3, full protection requires

$$s_{ij} = \beta.$$

As the expected return of sharing with many banks is the same as sharing with one other bank j, while the likelihood of being saved by the CB if required strictly increases. It follows that if a bank i shares its investments it will do so with one other bank and will consequently be fully protected.

□

3.3.4 Sharing rule

Lemma 3.3 establishes that to be fully protected banks share their portfolios with exactly one other bank and they will share the minimum level of capital that will force the CB to intervene, namely β. In addition, Corollary 3.2 establishes that if a high type bank chooses to share it's investment, it will ensure that it is fully protected.

Therefore, we can now determine under what conditions will banks choose to form a network. In particular, the expected returns of bank i when it invests all its capital to its own portfolio is:

$$\mathbb{E}[\Pi_i^{NS}] = pR_i \tag{3.3}$$

The return of bank i when the network is connected enough to ensure that the CB will provide liquidity to any distressed bank is:

$$\mathbb{E}[\Pi_i^{S}] = (1-\beta)R_i + \beta\mathbb{E}[R_j] \tag{3.4}$$

Since banks are risk neutral, bank i will choose not to form a network if the expected returns of doing so are higher, that is if:

$$
\begin{gathered}
\mathbb{E}[\Pi_h^{NS}] \geq \mathbb{E}[\Pi_h^{S}] \\
R^h \left(1 + \frac{(1-p)(\frac{1}{N}-1)}{\beta(1-\mu)}\right) \geq R^{l}
\end{gathered} \tag{3.5}
$$[1]

where Π_h^{NS} denotes the profits of a high type bank if it does not share its portfolio and Π_h^{S} denotes its profits if it does. Note, that a low type bank always wants to share its portfolio. Condition 3.5 is the main result of the benchmark model. It shows what determines whether banks will endogenously increase systemic risk. As one would expect, the probability of a shock affects negatively the decision not to share investment opportunities, whereas the bankruptcy threshold and the proportion of R^h, μ, affect it positively.

This is an intuitive result as the CB is assumed to have unlimited funds and has the ability to diminish risk completely from the perspective of banks. Moreover, the number of banks has a relatively small but positive effect on condition 3.5, as with even with a small number of banks the effect on the inequality becomes negligible.

3.4 Model with Depositors

3.4.1 Depositors

There are M depositors that have enough endowment to invest in all bank portfolios. They are born at the beginning of each period and live for 1 period. By assumption $M > N$. Depositors are risk neutral and have a utility that depends on their final period consumption. Thus, the utility of a depositor that is born in period t is:

$$
U(c_t) = c_{t+1}
$$

At their initial period, they invest their endowment in the banks, which offer a return d in the next period, a fraction of the total return that banks receive in each period. The outside option return to investment is set to d. This ensures that it is individually rational for consumers to deposit in any bank. In addition, portfolio returns are realized every period. In particular, each portfolio returns $R_i/2$ at $t = 1$ and $t = 2$ with probability p and 0 otherwise. These can be thought of as long-term portfolios

[1] For Derivation see Appendix 3.A.1.

as they yield returns for 2 periods but the positions held by banks cannot change in period 1. In addition, as, by assumption, there are more investors than banks, the bargaining power lies with banks, who offer the minimum amount possible. This being the case, bank types cannot be identified.

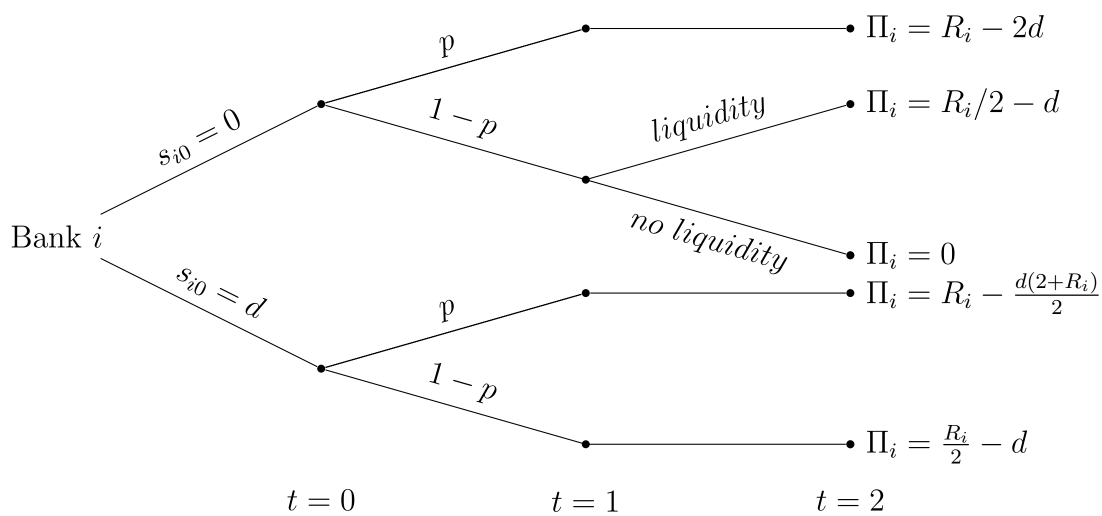
Thus,

$$0 < d < \frac{R^l}{2}$$

Hence, if a bank is unable to provide the pledged returns in period 1, it must liquidate its assets and gets zero returns in the next period. This endogenizes the event of a bankruptcy. It is beyond the purpose of this paper to analyse the liquidation process further. As before, a regulator can step in and provide the necessary liquidity in order for investors not to withdraw their capital.

We further extend the model to allow for banks to have access to a risk free asset, denoted by R_0, with return equal to 1, that banks can use to ensure they have the necessary liquidity to pay back depositors in period 1. We denote bank i's share of deposits invested in R_0 by s_{i0}.

Figure 3.4.1: Timeline with depositors and no sharing



This extension allows us to achieve two things. Firstly, bankruptcy is endogenised and corresponds to the situation that the banks do not have enough liquidity to satisfy their deposits. In addition, in this version, the CB provides enough liquidity for the depositors to be paid, but does not allow banks to keep excessive returns after their portfolios have suffered a loss which is a more realistic assumption on the type of bank interventions.

3.4.2 Sharing optimality with depositors

For any bank i the trade-off between sharing or not sharing depends on the amount d pledged to investors and the returns of other banks. In particular, an individual bank i will choose not to share its investment with other banks if its net returns are higher when it stores some the capital pledged to depositors in period 1.

Firstly, we assume that for any bank, in the absence of a CB, the returns from investing d in R_0 in period 0 and thus ensuring sustainability in period 2, are always higher than investing all of its available capital in period 0 and thus, going bankrupt in period 1 in case of a shock.

Assumption 3.1.

$$\mathbb{E}[\mathbf{R}_i^d] > \mathbb{E}[\mathbf{R}_i^{nd}]$$

$$p\left[R_i - \frac{d(2+R_i)}{2}\right] + (1-p)\left[\frac{R_i}{2} - d\right] > p[R_i - 2d]$$

$$2d(2p-1) > R_i(p(d+1)-1)$$

That is, the increase in profits that banks make if they ensure they will survive the second period are higher than the returns lost from storing instead of investing capital d when there is no shock.

As in the baseline model, banks share investments only in order to be saved by the CB in case of a crisis. In particular, banks that have projects with high return, will be unwilling to invest in a project that may have lower return, and do so only if the CB is going to intervene.

Note, that in order for any bank i to ensure that that the initial mechanism holds for ensuring that the CB will intervene and provide them with enough liquidity, they need to invest to another bank. In order for this behaviour to be optimal it must be a best response for all banks. As before, banks with low returns will always be strictly better off by investing in any other banks, as the expected return of the other banks is higher than the return of their own project. Therefore, it is sufficient to show that this is a best response for any bank i with a high return project.

In order for the CB to intervene it must be the case that the failure of bank i has made at least one more bank fail. In order for bank j to fail it is necessary that it does not have enough liquidity to return to its investors. Also, since banks are expected to depend on the CB to provide liquidity, they do not have any incentives to invest in R_0. Thus assuming that bank j has been hit by a crisis, it must be the case that:

$$\frac{R_i}{2}(s_{ii} + s_{ij}) \leq d$$

That is the period 1 returns of bank i when bank j has been hit by a shock must be less than d in order for a CB intervention to be possible. But since, bank j has had a crisis, this is equal to:

$$\begin{aligned} \frac{R_i}{2} s_{ii} &\leq d \\ s_{ii} &\leq \frac{d}{R_i} \\ \Rightarrow s_{ij} \geq 1 - s_{ii} &\geq 1 - \frac{d}{R_i} \end{aligned} \tag{3.6}$$

That is, condition 3.6 is a necessary condition for any CB intervention. By Lemma 3.3, as a high type bank would not want to share more than necessary, full protection requires:

$$s_{ii} = \frac{d}{R_i} \Rightarrow s_{ij} = 1 - \frac{d}{R_i} \tag{3.7}$$

In addition, if 3.6 does not hold for bank i then it knows that the CB will not intervene and thus it must always invest d in R_0 by Assumption 3.1.

Therefore, following a logic similar to that in the baseline model, a bank i will find it optimal to share its investments if:

$$\begin{aligned} \mathbb{E}[\Pi_h^{NS}] - \mathbb{E}[\Pi_h^{S}] &> 0 \\ \frac{R^h}{k} &> \frac{1 + p^N(1 - \frac{2d}{k})}{1 + p(1 - d)} \end{aligned} \tag{3.8}$$ [2]

where,

$$k \equiv \frac{d}{R^h}R^h + (1 - \frac{d}{R^h})\phi$$

is the one period expected return of a high type bank connected with a single other bank and

$$\phi \equiv \frac{(\alpha N - 1)R^h + (1-\alpha)R^l}{(N-1)}$$

is the two period expected return of a bank of unknown type that is connected with a high bank.

Note also that as $N \to \infty$, condition 3.8 becomes

$$\lim_{N\to\infty} \frac{R^h}{k} > \frac{1}{1 + p(1-d)} \tag{3.9}$$

which always holds since $R^h > k$. This is means for very large banking networks banks always prefer not to share their investments. This makes sense as for very large networks the CB will intervene with certainty and hence all banks will make 0 profits in period 1. This result is starkly different with the baseline model where N was not important in determining banks' behavior. The reason behind this is that in the extended model the CB's intervention is costly to banks as they lose part of their profits.

Condition 3.8 is the main result of the paper. As in the benchmark model, whether bank i will participate in the mechanism that forces the CB to intervene, depends positively on the probability of the shock and negatively on the spread of returns and the proportion of low return projects. The effect of d is also similar to β, i.e. it decreases the willingness of banks to share investments.

[2]For the derivation of condition 3.8 see Appendix 3.A.2.

3.5 Concluding Remarks

In this chapter we develop a stylized model between risk neutral banks and a regulator. In our model banks can share investments through the interbank market thus taking on a share of the other bank's exposure to risk. Without the presence and expected intervention of a regulator, banks with access to high return portfolios would not do this, as it reduces their expected profits. However, if the regulator is expected to intervene then banks may choose to share investment as this ensures that they cannot go bankrupt. Therefore, in our model banks may choose to create networks in order to become systematically important.

In the baseline model, we find that banks either connect in a circle, where each is connected with a single other bank or they do not share investments with any other bank. Which of the two depends negatively on the difference between the return of the high and low portfolios, on the ratio of banks of high portfolios and on the bankruptcy threshold. It depends positively on the likelihood of a single bank being distressed. The number of banks in the system plays a negligible role in their behaviour. Crucially, these results depend on a fairly extreme decision rule of the regulator, who commits to save any bank if and only if not doing so causes systemic collapse.

Our baseline model is extended by adding depositors and thus endogenising the bankruptcy threshold, adding a safe asset banks can invest in and allowing the regulator to save depositors only and not banks, thereby making liquidity provision costly to banks, a more realistic assumption. Qualitatively, the results of this extended model remain similar. One stark difference is that in the model with depositors the number of banks is very important in determining the equilibrium. In particular as the number of banks increases the incentive to share investment decreases rapidly. This is because in the extended model liquidity provision is costly for banks. Consequently, as the number of banks increases, the likelihood that at least one will face a liquidity shock goes to one and hence all banks loose period 1 profits with certainty. This suggests that from a policy perspective, it is preferable for bailouts to be costly to banks.

Future research could extend the model by having the regulator maximize a social welfare function, introduce a richer environment of portfolio returns and consider the case of risk averse depositors and banks.

Appendix

Appendix 3.A

3.A.1 Derivation of condition 3.5

Given that there are N banks, where μ is the proportion of banks with high return and $1-\mu$ the proportion of banks with low return, R^h and R^l. We know from lemma 3.3 that $s_{ij}=\beta$, hence, $s_{ii}=1-\beta$. We also know, for a high portfolio bank:

$$\mathbb{E}[\Pi_h] = s_{ii}R^h + (1-s_{ii})\frac{(\mu N-1)R^h + (1-\mu)NR^l}{N-1} \tag{3.10}$$

Hence,

$$\begin{aligned}
\mathbb{E}[\Pi_h^{NS}] &\geq \mathbb{E}[\Pi_h^{S}] \\
pR^h &\geq (1-\beta)R^h + \beta\frac{(\mu N-1)R^h + (1-\mu)NR^l}{N-1} \\
R^h\frac{(\beta-1-p)(N-1)-\beta(\mu N-1)}{N-1} &\geq \frac{\beta(1-\mu)N}{N-1}R^l \\
R^h\left(1+\frac{(1-p)(\frac{1}{N}-1)}{\beta(1-\mu)}\right) &\geq R^l
\end{aligned}$$

3.A.2 Derivation of condition 3.8

Similarly to the previous derivation, bank i will find it optimal not to invest in any bank j if:

$$
\begin{aligned}
\mathbb{E}[\Pi_h^{NS}] - \mathbb{E}[\Pi_h^{S}] &> 0 \\
p\left[R^h - \frac{d(2+R^h)}{2}\right] + (1-p)\left[\frac{R^h}{2} - d\right] - \Bigg[p^N\Big(\frac{d}{R^h}R_h + (1-\frac{d}{R^h})\phi - 2d\Big) & \\
+(1-p^N)\Bigg(\Bigg(\frac{d}{R^h}R^h + (1-\frac{d}{R^h})\phi\Bigg)/2 - d\Bigg)\Bigg] &> 0 \\
p\left[R^h - \frac{d(2+R^h)}{2}\right] + (1-p)\left[\frac{R^h}{2} - d\right] - \Bigg[p^N\Big(k - 2d\Big) & \\
+(1-p^N)\Big(\frac{k}{2} - d\Big)\Bigg] &> 0 \\
\frac{1}{2}\left(2dp^N - dpR^h - kp^N - k + pR^h + R^h\right) &> 0 \\
2dp^N - dpR^h - kp^N - k + pR^h + R^h &> 0 \\
R^h\Big(1 + p(1-d)\Big) - k(1+p^N) + 2dp^N &> 0 \\
R^h\Big(1 + p(1-d)\Big) &> k(1+p^N) - 2dp^N \\
\frac{R^h\Big(1 + p(1-d)\Big)}{k} &> 1 + p^N - \frac{2dp^N}{k} \\
\frac{R^h}{k} &> \frac{1 + p^N - p^N\frac{2d}{k}}{1 + p - pd}
\end{aligned}
$$

where,

$$\phi \equiv \frac{(\alpha N - 1)R^h + (1-\alpha)R^l}{(N-1)}$$

is the two period expected return of a bank of unknown type that is connected with a high bank and

$$k \equiv \frac{d}{R^h}R^h + (1 - \frac{d}{R^h})\phi$$

is the one period expected return of a high type bank i connected with a single other bank and with $s_{ii} = \frac{d}{R^h}$.

Chapter 4

Electricity Sharing Agreement with no Commitment and no Observability

4.1 Introduction

An almost ubiquitous assumption in economics is that economic agents have intertemporal consumption-smoothing preferences. Faced with an uncertain income stream, the need to save arises in order to maintain future consumption levels. This of course presupposes the ability to save. Famously, Kocherlakota (1996) explored the properties of risk sharing when there is lack of commitment under symmetric information and showed that it may lead to imperfect risk sharing for impatient enough agents. However, there can be situations in which the assumption of symmetric information needs to be relaxed also. This paper creates an algorithm to simulate an enforceable, feasible and optimal agreement between to risk averse parties, where there is both lack of commitment and information asymmetry.

We develop and characterise a sharing agreement in which two countries that produce electricity can improve on their individual utilities by sharing the electricity they each generate. Electricity generation technology is stochastic and can be high or low. The electricity generated by each country is observable only to itself. In addition, countries cannot commit to a sharing agreement. The justification for this, is that this environment would fit mostly developing countries, in which electricity institutions may be less developed. For example, each country, due to a change in political regime, or other reasons, may opt to nullify a contract whenever they choose to. Moreover, we compare the welfare gains between being part of this agreement and the autarky utility level. We find that the agreement can increase the expected lifetime utility of each party by $1\% - 5\%$ depending on the level for risk aversion σ and impatience δ. Further, we find that both σ and δ have a strong effect on the average consumption level per state. σ's being negative, while δ's positive. Finally, we find that σ has a particularly strong effect on the amount of time each party stays at the autarky utility level. These results can be used to quantify the opportunity cost of not setting a bilaterally-trusted independent authority that monitors and publicly reports electricity generation.

In the next section the relevant literature on risk sharing is discussed, in section 4.3 the model is described. Section 4.4 characterizes the dynamic programming form of the model and describes the algorithm used to generate our results. Section 4.5 presents and discusses the results and section 4.6 concludes.

4.2 Literature Review

In his seminal paper Kocherlakota (1996) introduces a model of risk sharing between two consumers in a complete information environment without commitment. They show that if players are patient enough, then there exists an SPNE in which they will be able to achieve the first best. When players are not sufficiently patient, then efficient allocations have the identical distributions of consumption between players, which are positively correlated with current and lagged income. The strength of the model explored in Kocherlakota (1996) is that it is a simple yet ubiquitous premise. The crucial difference between our model and Kocherlakota's paper is that in the latter aggregate income is constant each period and hence known to all agents. Our model is also closely related to Thomas and Worrall (1988), in which they also examine non-enforceable arrangements between parties. Consequently parties can exit the agreement at any point, meaning that any feasible arrangement must satisfy each party's participation constraint at any period. Similarly, (Ligon, Thomas, and Worrall, 2002) characterize an efficient and informal sharing mechanism between villagers that want to pool their perishable and stochastically generated income. Ligon, Thomas, and Worrall show that at the optimum consumption is correlated to per-period income. Differently from us, these are environments of no aggregate uncertainty, while in our paper there exists aggregate uncertainty and therefore information asymmetry between parties.

There have been many extensions to the basic no-commitment, no aggregate risk sharing framework, both on the commitment side as well as on environments with incomplete information. For example, Ábrahám and Laczó (2018) introduce a storage technology which allows consumers to store their income, either publicly or privately, as well as share their income in each period. They find that consumers never have an incentive to store privately, which is unobserved by the other party and has the same return as the public storage technology. They find that the return on storage is a key determinant in the time-path of consumption. In a similar framework, that maintains the no aggregate risk assumption, Koeppl (2007) introduce a costly punishment technology that agents can use when the other party decides to revert to autarky. They show this technology is sometimes used on the equilibrium path. Introducing this technology is not possible in our framework due to our assumed aggregate uncertainty.

In order to satisfy both individual rationality and incentive compatibility constraints our mechanism is based on the idea of future promises to entice whoever has the highest current income to share with the other party. This type of mechanism has been explored in the gift exchange and trust literature. For example, Abdulkadiroğlu and Bagwell (2013) study a repeated game in which privately informed players are willing to share some of their income only if they anticipate a favour in the future. In their case the size of the favor may decrease over time. A key requirement for the existence of equilibrium in their set up is that favours can be reciprocated relatively frequently. Möbius (2001) shows that even very infrequent favours can be bilaterally sustained, if players are connected in networks which allow for a connected player to reciprocate on one's behalf. In our set up, the frequency with which sharing takes place affects the gains from trade significantly.

Our paper is most closely related to Hertel (2004) and Nikandrova and Steinbuks (2017). Hertel (2004) characterize a model which is like ours, with the difference that uncertainty is one sided. In other words, in that paper there is a risk neutral consumer with a steady income stream and one whose period income is privately known to be either high or low. This model does not fit our envisaged application where uncertainty must be two-sided. On the other hand, Nikandrova and Steinbuks (2017) consider a very similar application to ours, i.e. electricity sharing between two countries and therefore make many similar assumptions in terms of the interaction. However, instead of numerically approximating the optimal sharing contract, they consider a sub-optimal token sharing mechanism that can be fully characterized.

4.3 Model

The model constructed in this paper extends the Efficient and Sustainable Risk Sharing With Adverse Selection model, as developed in Hertel, 2004. In particular we extend that framework by adding uncertain income generation for both parties. Our model consists of two risk averse countries who engage in risk sharing of a non-storable good, over an infinite horizon. Each country's generated good is stochastic. This section describes the model before the dynamic form of the sharing agreement is stated. The numerical results are then presented and discussed.

There are two countries, henceforth denoted by $i \in \{1, 2\}$. Each country owns an electricity generation technology. Electricity generated each period can be high with probability q for country 1 and with probability p for country 2. We assume that the levels of electricity generated by each country are independent. Hence, in each time period there are four possible states of the world denoted by $S = \{(s_1, s_2) : s_i \in S_i \text{ for } i \in \{1, 2\}\}$, with $S_i = \{L, H\}\ \forall\, i$. Clearly, the first position of each element in S, denotes the level of electricity produced by country 1 and the second that by country 2. In addition, $s_i^j \in S_i$ refers to the state of player i independently to the state of the other player, with $j \in \{L, H\}$. As a result of our assumption on independence between states, the probability of each state is $p(s) \in \{(1-q)(1-p), (1-q)p, q(1-p), qp\}$. For readability, we denote the amount electricity generated in state s^j by each country, i.e. $e^{s_i^j}$, by e_i^j and the aggregate electricity generated in each state by e^s. That is, e_1^h refers the electricity generated by country 1 being high and e^{hh} to the aggregate electricity generated in state s^{hh} . Hence, the possible aggregate levels of electricity for every time period are:

$$e^s = \begin{cases} e^{hh} = e_1^h + e_2^h, \\ e^{hl} = e_1^h + e_2^l, \\ e^{lh} = e_1^l + e_2^h, \\ e^{ll} = e_1^l + e_2^l, \end{cases}$$

In addition, the states between time periods are assumed to be identically and independently distributed, with s_t^j being the state of the world in period t. A sequence $(s_1^j, s_2^j, \ldots, s_t^j)$ is a history of states denoted by h_t. $\mathcal{H}_t := \{LL, LH, HL, HH\}^t$ denotes the set of all histories in period t and $\mathcal{H} = \bigcup_{t=1}^{\infty} \mathcal{H}_t \cup \{h_0\}$ that of all possible histories. It follows by the independence of states that the probability of each history h_t is,

$$p(h_t) = p^n(1-p)^{t-n}q^m(1-q)^{t-m}$$

where n is the number of high states in t periods for country 2 and m is the equivalent for country 1.

A mapping $c : \mathcal{H} \to \mathbb{R}$ is a sharing rule that describes the consumption of electricity for country 1 for every $h_t \in \mathcal{H}$. We denote by c_{h_t} the consumption of electricity by country 1 after any history h_t according to the sharing rule c. In addition, $c^s_{h_{t-1}}$ is the consumption of country 1 under the sharing rule c, with history h_{t-1} and s_t. It follows that the electricity consumption of country 2 is $e^s_t - c^s_{h_{t-1}}$, where e^s_t is the aggregate electricity generated in period t. The sharing rule c^* is the autarky sharing rule, in other words, the sharing rule in which no transfers are made and each country consumes all the electricity it generates at every period.

Countries' periodic utilities u_i are strictly concave and increasing. They exhibit non-increasing absolute risk aversion. That is, for x, y and $z \in \mathbb{R}$, $\alpha > 0$ and $\lambda \in (0,1)$,

$$\lambda \upsilon_i(x) + (1-\lambda)\upsilon_i(y) \geq \upsilon_i(z)$$

implies

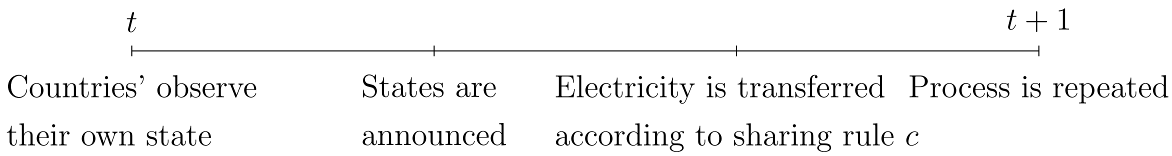
$$\lambda \upsilon_i(x+\alpha) + (1-\lambda)\upsilon_i(y+\alpha) \geq \upsilon_i(z+\alpha),$$

for $i = \{1, 2\}$. Therefore the expected utility for country 1 under any sharing rule c that induces true revelation of the state by both countries is:

$$U^i(c) = \sum_{t=1}^{\infty} \delta^{t-1} \sum_{h_t \in \mathcal{H}} u(c_{h_t})p(h_t)$$

Given the above, the game is as follows. Each period, the countries observe their own production of electricity and not the electricity produced by the other country. They each announce their own levels of production. The sharing rule c specifies how much electricity will be transferred between each other. Electricity is consumed instantaneously and the next period the process starts from the beginning. The process can be seen in figure 1.

Figure 4.3.1: Within period timeline



We aim to construct a unique sharing rule that is pareto dominant, efficient and sustainable. In other words, we want to describe a sharing rule c such that both countries' would be willing to participate in and that there exists no other sharing rule that would be pareto dominating, in the sense that it improves the lifetime utility of any of the two players without making the other worse off. In the following section we describe the necessary conditions of such a sharing rule.

4.3.1 Feasibility and sustainability

As stated above, electricity is considered a non-storable good. As such, feasibility of any sharing rule dictates that the total consumption of electricity in any period t must not exceed the total generation at that period. Clearly, given that countries' utility is strictly increasing in electricity consumption, all electricity generated in any period will be consumed.

$$0 \leq c^j_{h_{t-1}} \leq e^j_t, \quad \forall \;\; t \tag{F}$$

The participation constraint states that any sharing rule must offer a weakly higher expected utility to both countries than the expected utility that they obtain if they deviate to the autarky sharing rule, c^*. The rationale behind this constraint is straightforward. Since, countries cannot commit to any sharing level prior to the revelation of each state, they will trade electricity if, and only if, in doing so they achieve a weakly higher expected utility than their expected utility when no sharing takes place. That is, the participation constraint dictates:

$$U^i_H(c_{h_{t-1}}) \geq u^i(e^h_i) + \delta U^i_s(c^*), \tag{P^i_H}$$
$$U^i_L(c_{h_{t-1}}) \geq u^i(e^l_i) + \delta U^i_s(c^*), \tag{P^i_L}$$

for all $h_{t-1} \in \mathcal{H}$ and countries 1 and 2.

$U^i_{H(L)}(c_{h_{t-1}})$, refers to the expected utility of country i under sharing rule c and h_{t-1}, when their individual state at t is $h(l)$ but they do not know the amount of electricity generated by the other country.

In addition, given that there is no commitment, P^i_L sets the lower limit of the expected utility that any sharing agreement must satisfy. Therefore we will only consider sharing rules that guarantee to each country an expected utility equal to their utility after s^i_l and the continuation autarky utility, that is, $U^i(c) \geq u^i(e^i_l) + \delta U^i_s(c^*)$ for all t and all $h_t \in \mathcal{H}$. More generally, as the states are independent and electricity is non-storable, the interim utility, i.e. the utility for each country, after state s is realised and before they know the state of the other country, is:

$$U^i_{s^j_i}(c) = u_i(c(s^j_i)) + \delta U^i(c(s^j_i))$$

Where $u_1(c(s^j_1))$ refers to the expected utility of country 1 in state s^j_1 and history h_t. Similarly, $u_2(c(s^j_2))$, is the expected utility of country 2, when her state is s^j_2.

The revelation constraint, R, requires that both countries must find it weakly optimal to announce their state truthfully, for all h_{t-1} and all s^j_i. That is:

$$U^i_H(c_{h_{t-1}}) \geq u^i(c^l_{h_{t-1}} + e^h_i - e^l_i) + \delta U^i(c^l_{h_{t-1}}) \tag{R^i_H}$$
$$U^i_L(c_{h_{t-1}}) \geq u^i(c^h_{h_{t-1}} - e^h_i + e^l_i) + \delta U^i(c^h_{h_{t-1}}) \tag{R^i_L}$$

Note that R^i_L will always be satisfied if both $c^{hh}_{h_{t-1}} - (e^h_1 - e^l_1) < 0$ and $c^{hl}_{h_{t-1}} - (e^h_1 - e^l_1) < 0$, as this makes the consumption of country 1 in period t negative. Similarly R^2_L will always be satisfied if, $e^{hh} - c^{hh}_{h_{t-1}} - (e^h_2 - e^l_2) < 0$ and $e^{ll} - c^{lh}_{h_{t-1}} - (e^h_2 - e^l_2) < 0$.

Henceforth, we will assume that this is always the case and ignore these constraints. Indeed, this will be violated when the distribution of the two levels of electricity is very concentrated, thus decreasing the relevance of our model as the need for them to share electricity decreases substantially.

A sharing rule c is sustainable if it is a subgame perfect Nash equilibrium (SPNE) outcome of the agreement. That is, if there is no profitable deviation for any player in any period. Hence, any sharing agreement is a SPNE if the feasibility, participation and revelation constraints are satisfied for all $h_{t-1} \in \mathcal{H}$.

Definition 4.1. *A sustainable sharing rule is a function $c : \mathcal{H} \to \mathbb{R}$ that satisfies P, R and F for all $h_{t-1} \in \mathcal{H}$.*

4.3.2 Efficiency

As discussed above, the constraints R_L^i and P_L^i will always be satisfied, according to our assumptions. Hence we define the set of sustainable sharing rules as:

$$\mathcal{C} = \{c | c : \mathcal{H} \to \mathbb{R} \textit{ and } c \textit{ satisfies } P_H^i, R_H^i \textit{ and } F \ \ \forall h_{t-1} \in \mathcal{H}\}$$

We are interested in the sharing rules within $\mathcal{C}$, which are efficient, in the sense that they are not Pareto dominated. To find these we first characterize the actions taken in equilibrium by both countries.

Given a sharing rule c, at any history $h_t = (h_{t-1}, s_t^j)$ countries 1 and 2 announce their states truthfully. Country 1 makes a transfer $\max\{e_1^j - c_{h_{t-1}}^j, 0\}$ and country 2 makes a transfer $\max\{c_{h_{t-1}}^j - e_1^j, 0\}$. Deviation from any player triggers a permanent reversion to the autarky contract, c^*, i.e. zero transfers from both countries for all h_t. This strategy profile induces a sharing rule, that is a subgame perfect nash equilibrium of the sharing model. As long as R_H^i and P_H^i are satisfied, no country has an incentive to deviate from their strategy. This leads to the following definition.

Definition 4.2. *A sharing rule c is efficient in the set $\mathcal{C}$ if $c \in \mathcal{C}$ and c is not Pareto dominated by any other $c' \in \mathcal{C}$. The set of efficient sharing rules in $\mathcal{C}$ is denoted by $\mathcal{E}(\mathcal{C})$.*

4.3.3 Formulation of maximization problem

The following maximization problem characterizes a unique, efficient and sustainable sharing rule.

$$W(k) = \max\ U^2(c) \text{ subject to } c \in \mathcal{C} \text{ and } U^1(c) \geq k.$$

Let, $S(k)$ be the set of sharing rules that solve the above maximization problem. Hence,

$$S(k) = \{c | c \in \mathcal{C} \text{ and } U^1(c) \geq k \text{ and } U^2(c) = W(k)\}$$

Then the set of country 1 attainable utilities is:

$$\mathcal{K} = \{k | \ \exists\ c \in \mathcal{C} \text{ s.t. } U^1(c) = k\}$$

Therefore, according to this maximization formulation, country 2 maximizes its discounted lifetime expected utility subject to the discounted lifetime expected utility of country 1 being equal to k and P, R, F being satisfied for both countries. We use Proposition 1 and Lemma 1 to show existence and uniqueness of the proposed efficient sharing rule.

First, we show that although $\mathcal{C}$ is not a convex set, we can always construct Pareto dominating combinations of sustainable sharing rules from $c', c'' \in C$ that are part of C. Formally:

Lemma 4.1. *For any $c', c'' \in \mathcal{C}$ such that $c' \neq c''$ and any $\lambda \in (0, 1)$, there exists $c \in \mathcal{C}$ such that*

$$\begin{aligned} U^1(c') &= \lambda U^1(c') + (1 - \lambda)U^1(c''), \\ U^2(c') &\geq \lambda U^2(c') + (1 - \lambda)U^2(c''). \end{aligned} \tag{4.1}$$

Proof. See Appendix. □

Lemma 4.1 shows that although the set $\mathcal{C}$ may not be convex, for any two different sharing rules $\in \mathcal{C}$ we can always construct an sharing rule that is also $\in \mathcal{C}$ and in which both countries are at least as well of. This allows to use the following proposition.

Proposition 4.1. *$S(k)$ exists when $\mathcal{K} = [k^*, k^{**}]$ is non empty, has a single element for each $k \in \mathcal{K}$ and $W : \mathcal{K} \to \mathbb{R}$ is a concave and strictly decreasing function. In* *addition,* $k^* = \min \mathcal{K} = U^1(c^*)$, $k^{**} = \sup \mathcal{K}$ *and* $W(k^{**}) = U^2(c^*)$.

Proof. See Hertel (2004) Proposition 1. $\square$

It follows directly from Proposition 4.1 that the maximization problem described above characterizes a unique and efficient sharing rule, which exists as long as $\mathcal{K}$ is non-empty.

4.4 Computation

4.4.1 Dynamic Programming form

In this section we state $W(k)$ in a dynamic programming form. This will enable us to get numerical solutions for c^k, henceforth the sharing agreement that is a solution to $S(k)$.

By proposition 4.1 the sharing agreement maximizes the utility of country 2:[1]

$$\begin{aligned} W(k) = \ &\max qp[\upsilon_2(e_{hh} - c_{hh}) + \delta W(C_{hh})] + q(1-p)[\upsilon_2(e_{hl} - c_{hl}) + \delta W(C_{hl})] \\ &+ (1-q)p[\upsilon_2(e_{lh} - c_{lh}) + \delta W(C_{lh})] + (1-q)(1-p)[\upsilon_2(e_{ll} - c_{ll}) + \delta W(C_{ll})] \end{aligned} \tag{4.2}$$

subject to the promise made to country 1, k:

$$\begin{aligned} &\text{s.t. } qp[\upsilon_1(c_{hh}) + \delta C_{hh}] + q(1-p)[\upsilon_1(c_{hl}) + \delta C_{hl}] \\ &+ (1-q)p[\upsilon_1(c_{lh}) + \delta C_{lh}] + (1-q)(1-p)[\upsilon_1(c_{ll}) + \delta C_{ll}] \geq k \end{aligned} \tag{4.3}$$

In addition, in order to ensure truth telling by both countries after any promise the revelation constraints of both countries' when electricity generation is high must be satisfied:

[1]For clarification on notation used here see Appendix 4.A.2.

$$p[\upsilon_1(c_{hh}) + \delta C_{hh}] + (1-p)[\upsilon_1(c_{hl}) + \delta C_{hl}] \geq$$
$$p[\upsilon_1(c_{lh} + (e_{hh} - e_{lh})) + \delta C_{lh}] + (1-p)[\upsilon_1(c_{ll} + (e_{hl} - e_{ll})) + \delta C_{ll}], \qquad (R^1_H)$$
$$q[\upsilon_2(e_{hh} - c_{hh}) + \delta W(C_{hh})] + (1-q)[\upsilon_2(e_{lh} - c_{lh}) + \delta W(C_{lh})]$$
$$\geq q[\upsilon_2(e_{hl} - c_{hl} + (e_{hh} - e_{lh})) + \delta W(C_{hl})] + (1-q)[\upsilon_2(e_{ll} - c_{ll} + (e_{lh} - e_{ll})) + \delta W(C_{ll})] \quad (R^2_H)$$

Further, as there is no commitment, any country can walk away from the agreement at any period. To ensure that this does not happen, the participation constraints have to be satisfied when electricity generation is high:

$$q[\upsilon_2(e_{hh} - c_{hh}) + \delta W(C_{hh})] + (1-q)[\upsilon_2(e_{lh} - c_{lh}) + \delta W(C_{lh})] \geq U^2(c^*), \qquad (P^2_H)$$
$$p[\upsilon_1(c_{hh}) + \delta C_{hh}] + (1-p)[\upsilon_1(c_{hl}) + \delta C_{hl}] \geq U^1(c^*) \qquad (P^1_H)$$

Finally, feasibility requires that:

$$(c_{ll}, c_{hl}, c_{lh}, c_{hh}, C_{ll}, C_{hl}, C_{lh}, C_{hh}) \in [0, e_{ll}] \times [0, e_{hl}] \times [0, e_{lh}] \times [0, e_{hh}] \times \mathcal{K}^4, \qquad (4.4)$$

where the functions, $c_{ij} : \mathcal{K} \to [0, e_{ij}]$ and $C_{ij} : \mathcal{K} \to \mathcal{K}$ iteratively characterise $c^k \; \forall \; k \in \mathcal{K}$. Therefore, the sharing rule is characterized by the tuple $\Big(c_{ll},\, c_{lh},\, c_{hl},\, c_{hh},\, C_{ll},\, C_{lh},\, C_{hl},\, C_{hh}\Big)$: at h_0 $U^1(c^k) = k$. In the next period, consumption is determined by c_{ij} and k changes according to C_{ij}.

4.4.2 Description of algorithm

This section briefly explains how the algorithm works.[2] Firstly, we create a grid from G^* to G^{**}, with $G^* = U^1_{aut}$.[3] G^{**} is initially unknown but it must be the case that $VF(G^{**}) = U^2_{aut}$. This is because, the value function is strictly decreasing in G by proposition 4.1 and the lowest utility that will satisfy player 2's participation constraint is that which gives her the same utility as her autarky utility. The algorithm initially guesses G^{**} randomly. Matrix BG is created which contains all possible permutations

[2]For the full Mathematica code see appendix 4.A.5

[3]The notation used in this section is the one used in the algorithm. For a legend see appendix 4.A.4.

of four values selected from the grid, G. Each row in BG represents future promises to player 1, one for each of the four possible states at time t. That is, the first element of each row in BG is a candidate for a promise utility to player 1 after the state ll is announced, the second after the state lh, the third after hl and the last after hh.

The next step is to create matrix $kBG12c$. This is a matrix with 9 columns. The first is a value taken from grid, G, the next four are four possible levels of future promises, i.e. one for each state. The final four columns denote the position of these in grid G. The number of rows of $kBG12c$ is equal to the number of rows in BG times the number of rows in G. That is, the matrix $kBG12c$ contains all possible combinations of promises after each state realization in the next period, for all grid points of today's promised utility, G_t.

Using matrix $kBG12c$ and a random vector, named $Vsgsg$, that has the same number of elements as G and is an initial guess for the value corresponding to each promise level in G, arguments cll, clh, chl and chh are estimated. These values are consumption levels for player 2, that maximize player 2's utility for all promise levels and all combinations of future state contingent promises, given our guessed values. An augmented matrix is created, named $KbgBG12$. It has the same number of rows as $kBG12c$ but has additional columns for the estimated optimal consumption levels for each element in G. $KbgBG12$ is then partitioned into submatrices, each with the same number of columns and each row's first element being the same level value. That is, each submatrix corresponds to all the possible future promises and current levels of consumption for the same level of promised utility. For each grid value, the highest utility that player 2 can achieve, given that all the constraints are satisfied, is then used as the guess of the value associated with that grid value.

This process is repeated, with initial guesses the output of the previous guesses until the value function converges, i.e. the output grid and the input grid are the same up to some level of accuracy. The final step is finding the upper bound of G. That is, the value of G^{**} such that $VF(G^{**}) = U^2_{aut}$. This involves a second level of iteration. If this condition doesn't hold we make a new guess for G^{**} and repeat the entire process. We repeat the whole process until $VF(G^{**}) = U^2_{aut}$.

4.5 Results

The purpose of this paper is to assess the extent to which two-sided uncertainty hampers the scope for agents to trade. This is particularly important in the example of electricity sharing as it indicates the extent to which an external independent observer is necessary. In addition, we are also interested to qualitatively assess the relative importance the risk aversion and patience of agents' have on the long-term consumption properties of these agreements. To this end we use the algorithm to estimate the long-term utility of agents and compare to the autarky case for different values of risk aversion and and patience coefficients.

Our numerical results are derived using the algorithm that exploits the dynamic formulation of the maximization problem. We assume that consumers have a CRRA utility function of the form $u(c) = \frac{c^{(1-\sigma)}}{1-\sigma}$. Also, $e_l^1 = e_l^2 = 0.1$ and $e_h^1 = e_h^2 = 1.9$. In addition, we assume that the risk aversion coefficient, σ and the impatience coefficient, δ, are the same for both countries. Moreover, $p = q = 0.5$ Under these assumptions, the following figures show the estimated value function as well as the estimated functions for $c_{ij}(k)$ and $C_{ij}(k)$, for all values of k.

Table 4.5.1 shows the expected consumption levels for each country based on the announced states. These can be compared with the level of 0.1 and 1.9 that they are able to consume on their own when their individual state is low and high respectively. As can be seen by the table, the sharing agreement makes them consume less than their income sometimes when both countries are in the bad state. This is because their period consumption depends both on the current state and on the promised utility $\mathcal{K}$. Consequently, if an agent's promise to the other is high, they may need to consume less than their income even in state l.

In addition one can see both the effect of σ and on δ in table 4.5.1. In particular, as σ increases from 0.3 to 0.8, consumption levels decrease by approximately 20% when $\delta = 0.9$ and by approximately 10.8% when $\delta = 0.96$. Both these results are intuitive. The fact that as risk aversion increases period consumption decreases is stemming from the fact that more risk averse agents have stronger concern for consumption smoothing and as such, care relatively more about changes in their promised level of utility $\mathcal{K}$. Further, higher values of impatience, δ, imply the opposite, agents care about today's consumption more and hence the effect of σ is relatively smaller for high δ.

Table 4.5.1: Average Consumption

Agents' average consumption per state - $c_{\mathrm{ll}}, c_{\mathrm{lh}}, c_{\mathrm{hl}}, c_{\mathrm{hh}}$		
	$\delta = 0.9$	$\delta = 0.96$
$\sigma = 0.3$	0.1041, 0.5535, 1.4597, 1.8620	0.0921, 0.6104, 1.2817, 1.9011
$\sigma = 0.8$	0.0649, 0.4119, 1.2049, 1.8692	0.0776, 0.6769, 0.9571, 1.6525

Table 4.5.2 on the other shows the percentage time each country's expected utility is equal to the autarky utility under the agreement. Two observations are striking from this table. First, being part of the agreement means that the vast majority of time countries enjoy an expected lifetime utility higher then the autarky utility. Second, the degree of risk aversion has a very significant impact on the amount of time spent in the autarky utility level. The magnitude of the second effect is fairly similar for both values of δ. One hypothesis about this result is that more risk averse consumers are willing to give up today's consumption for higher future promise. But this means that a bad streak of electricity generation can relatively easily send someone to the autarky utility level. Put differently, more risk averse consumers are more sensitive to changes in promised utility, which makes it more variable and thus more likely to reach the autarky level.

Table 4.5.2: Time in Autarky

Percentage time spent in autarky utility		
	$\delta = 0.9$	$\delta = 0.96$
$\sigma = 0.3$	8.7%	9.2%
$\sigma = 0.8$	18.6%	20.8%

Table 4.5.3 shows the difference in expected lifetime utility between the autarky and the sharing agreement. It can be seen that the improvement, depending on the level of risk aversion and patience, varies $1\% - 5\%$ approximately. In this case it seems the more significant effect is that of the patience parameter δ. The more patient a country is, i.e. the more it values it's future consumption, the less willing it is to trade future earning for consumption smoothing.

Table 4.5.3: Lifetime Utility

Expected Lifetime Utility (Autarky)		
	$\delta = 0.9$	$\delta = 0.96$
$\sigma = 0.3$	13.17 (12.62)	32.09 (31.55)
$\sigma = 0.8$	45.81 (44.20)	111.35 (110.5)

4.6 Concluding Remarks

In this paper we study a dynamic sharing agreement without commitment and with asymmetric information between two risk averse parties who each generate a perishable consumption good stochastically. We describe an efficient, sustainable and feasible sharing rule that allows the parties to share the good between them by satisfying their incentive compatibility and participation constraints in each period and any history. The agreement is based on the idea of a promised utility level that one party owes to the other and that has to be satisfied each period. The level of that promised utility varies each period according to each party's announcement and the previous period's promise.

We construct an algorithm to simulate the agreement and study it's long-term properties. We find that the agreement can increase the expected lifetime utility of each party by $1\% - 5\%$ depending on the level for risk aversion σ and impatience δ. Further, we find that both σ and δ have a strong effect on the average consumption level per state. σ's being negative, while δ's positive. Finally, we find that σ has a particularly strong effect on the time each party stays at the autarky utility level. The intuition behind this is that a high degree of risk aversion makes parties' utility more sensitive to changes in promised utilities, which causes them to fluctuate more and thus spend more time on the autarky level also.

In the context of electricity sharing between two developing countries, these results can be used to quantify the opportunity cost of not setting a bilaterally-trusted independent authority that monitors and publicly reports electricity generation. As our results are relatively crude estimates due to the coarseness of our grid as well as the few parameter values used for calibration, further development of our algorithm would be a natural next step in order to provide more accurate results.

Appendix

Appendix 4.A

4.A.1 Proof of Lemma 4.1

Proof. To construct c from c' and c'' choose $c(h_t)$ s.t.

$$\upsilon_1(c(h_t)) = \lambda\upsilon_1(c'(h_t)) + (1-\lambda)\upsilon_1(c''(h_t)), \qquad \forall h_t \in \mathcal{H}$$

As υ_1 is strictly concave, $c(h_t) \leq \lambda c'(h_t) + (1-\lambda)c''(h_t)$ hence:

$$\upsilon_2(e(h_t) - c(h_t)) \geq \lambda\upsilon_2(e(h_t) - c'(h_t)) + (1-\lambda)\upsilon_2(e(h_t) - c''(h_t)), \qquad \forall h_t \in \mathcal{H}$$

Hence, c satisfies Lemma 1 and P_H^2. We therefore need to show that it also satisfies P_H^1, R_H^2 and R_H^1, so that $c \in C$ indeed.

First, P_H^2 will definitely be satisfied as by construction $U^2(c)$ is greater than both $U^2(c')$ and $U^2(c'')$ and therefore it must be greater than $U^2(c_h^*)$. A similar argument holds for P_H^1. In particular since $U^1(c) = \lambda U^1(c') + (1-\lambda)U^1(c'')$ it must be the case that, $U^1(c) > \min\{U^1(c'), U^1(c'')\}$, when $c' \neq c''$. As both sharing rules c' and c'' satisfy P_H^1, so does c.

Finally, we show that R_H^1 is also satisfied. By assumption:

$$p[\upsilon_1(c'_{hh}) + \delta C'_{hh}] + (1-p)[\upsilon_1(c'_{hl}) + \delta C'_{hl}] \geq p[\upsilon_1(c'_{lh} + (e_{hh} - e_{lh})) + \delta C'_{lh}] + (1-p)[\upsilon_1(c'_{ll} + (e_{hl} - e_{ll})) + \delta C'_{ll}]$$
$$p[\upsilon_1(c''_{hh}) + \delta C''_{hh}] + (1-p)[\upsilon_1(c''_{hl}) + \delta C''_{hl}] \geq p[\upsilon_1(c''_{lh} + (e_{hh} - e_{lh})) + \delta C''_{lh}] + (1-p)[\upsilon_1(c''_{ll} + (e_{hl} - e_{ll})) + \delta C''_{ll}]$$

Then,

$$
\begin{aligned}
&p[\upsilon_1(c_{hh}) + \delta C_{hh}] + (1-p)[\upsilon_1(c_{hl}) + \delta C_{hl}] = \\
&\lambda\Big[p[\upsilon_1(c'_{hh}) + \delta C'_{hh}] + (1-p)[\upsilon_1(c'_{hl}) + \delta C'_{hl}]\Big] + (1-\lambda)\Big[p[\upsilon_1(c''_{hh}) + \delta C''_{hh}] + (1-p)[\upsilon_1(c''_{hl}) + \delta C''_{hl}]\Big] \geq \\
&\lambda\Big[p[\upsilon_1(c'_{lh} + (e_{hh} - e_{lh})) + \delta C'_{lh}] + (1-p)[\upsilon_1(c'_{ll} + (e_{hl} - e_{ll})) + \delta C'_{ll}]\Big] + \\
&(1-\lambda)\Big[p[\upsilon_1(c''_{lh} + (e_{hh} - e_{lh})) + \delta C''_{lh}] + (1-p)[\upsilon_1(c''_{ll} + (e_{hl} - e_{ll})) + \delta C''_{ll}]\Big] = \\
&\lambda\Big[p[\upsilon_1(c'_{lh} + (e_{hh} - e_{lh})) + (1-p)[\upsilon_1(c'_{ll} + (e_{hl} - e_{ll}))\Big] \\
&+ (1-\lambda)\Big[p[\upsilon_1(c''_{lh} + (e_{hh} - e_{lh})) + (1-p)[\upsilon_1(c''_{ll} + (e_{hl} - e_{ll}))\Big] + \delta(p\, C_{lh} + (1-p)\, C_{ll})
\end{aligned}
$$

Note that:

$$
\begin{aligned}
&\lambda\Big[p[\upsilon_1(c'_{lh} + (e_{hh} - e_{lh})) \\
&+ (1-p)[\upsilon_1(c'_{ll} + (e_{hl} - e_{ll}))\Big] + (1-\lambda)\Big[p[\upsilon_1(c''_{lh} + (e_{hh} - e_{lh})) + (1-p)[\upsilon_1(c''_{ll} + (e_{hl} - e_{ll}))\Big] \geq \\
&p[\upsilon_1(c_{lh} + (e_{hh} - e_{lh}))] + (1-p)[\upsilon_1(c_{ll} + (e_{hl} - e_{ll}))]
\end{aligned}
$$

Since, υ is a strictly concave function and by construction $U^1(c) = \lambda U^1(c') + (1-\lambda)U^1(c'')$. Hence, the above inequality must be true since a constant positive term is added to all the consumption levels of country 1.

It follows that:

$$
\begin{aligned}
&p[\upsilon_1(c_{hh}) + \delta C_{hh}] + (1-p)[\upsilon_1(c_{hl}) + \delta C_{hl}] \geq \\
&p[\upsilon_1(c_{lh} + (e_{hh} - e_{lh})) + \delta C_{lh}] + (1-p)[\upsilon_1(c_{ll} + (e_{hl} - e_{ll})) + \delta C_{ll}
\end{aligned}
$$

□

4.A.2 Definitions for arguments in $W(k)$

Define, $\forall k \in K$:

$$
\begin{aligned}
c_{ll}(k) &= c^k(LL) \rightarrow \text{player 1 consumption at } s_{ll} \text{ s.t. } k \\
c_{lh}(k) &= c^k(LH) \rightarrow \text{player 1 consumption at } s_{lh} \text{ s.t. } k \\
c_{hl}(k) &= c^k(HL) \rightarrow \text{player 1 consumption at } s_{hl} \text{ s.t. } k \\
c_{hh}(k) &= c^k(HH) \rightarrow \text{player 1 consumption at } s_{hh} \text{ s.t. } k \\
C_{ll}(k) &= U^1(c^k_{LL}) \rightarrow \text{player 1 expected lifetime utility after state } s_{ll} \text{ s.t. } k \\
C_{lh}(k) &= U^1(c^k_{LH}) \rightarrow \text{player 1 expected lifetime utility after state } s_{lh} \text{ s.t. } k \\
C_{hl}(k) &= U^1(c^k_{HL}) \rightarrow \text{player 1 expected lifetime utility after state } s_{hl} \text{ s.t. } k \\
C_{hh}(k) &= U^1(c^k_{HH}) \rightarrow \text{player 1 expected lifetime utility after state } s_{hh} \text{ s.t. } k
\end{aligned}
$$

4.A.3 Algorithm

$$
\begin{aligned}
W_n(k_n) = \ & \max\ qp[\upsilon_2(3.8 - c_{hh}) + \delta W_{n-1}(C_{hh})] + q(1-p)[\upsilon_2(2 - c_{hl}) \\
& + \delta W_{n-1}(C_{hl})] + (1-q)p[\upsilon_2(2 - c_{lh}) + \delta W_{n-1}(C_{lh})] + (1-q)(1-p)[\upsilon_2(0.2 - c_{ll}) + \delta W_{n-1}(C_{ll})] \\
& \text{s.t. } qp[\upsilon_1(c_{hh}) + \delta C_{hh}] + q(1-p)[\upsilon_1(c_{hl}) + \delta C_{hl}] \\
& + (1-q)p[\upsilon_1(c_{lh}) + \delta C_{lh}] + (1-q)(1-p)[\upsilon_1(c_{ll}) + \delta C_{ll}] \geq k_n, \\
& p[\upsilon_1(c_{hh}) + \delta C_{hh}] + (1-p)[\upsilon_1(c_{hl}) + \delta C_{hl}] \geq p[\upsilon_1(c_{lh} + 1.8) + \delta C_{lh}] + (1-p)[\upsilon_1(c_{ll} + 1.8) + \delta C_{ll}], && (R^1_H) \\
& q[\upsilon_2(3.8 - c_{hh}) + \delta W_{n-1}(C_{hh})] + (1-q)[\upsilon_2(2 - c_{lh}) + \delta W_{n-1}(C_{lh})] \geq q[\upsilon_2(2 - c_{hl} + 1.8) + \delta W_{n-1}(C_{hl})] \\
& + (1-q)[\upsilon_2(0.2 - c_{ll} + 1.8) + \delta W_{n-1}(C_{ll})], && (R^2_H) \\
& q[\upsilon_2(3.8 - c_{hh}) + \delta W_{n-1}(C_{hh})] + (1-q)[\upsilon_2(2 - c_{lh}) + \delta W_{n-1}(C_{lh})] \geq U^2(c^*), && (P^2_H) \\
& q[\upsilon_2(2 - c_{hl}) + \delta W_{n-1}(C_{hl})] + (1-q)[\upsilon_2(0.2 - c_{ll}) + \delta W_{n-1}(C_{ll})] \geq U^2(c^*), && (P^2_L) \\
& p[\upsilon_1(c_{hh}) + \delta C_{hh}] + (1-p)[\upsilon_1(c_{hl}) + \delta C_{hl}] \geq U^1(c^*), && (P^1_H) \\
& p[\upsilon_1(c_{lh}) + \delta C_{lh}] + (1-p)[\upsilon_1(c_{ll}) + \delta C_{ll}] \geq U^1(c^*), && (P^1_L) \\
& (c_{ll}, c_{hl}, c_{lh}, c_{hh}, C_{ll}, C_{hl}, C_{lh}, C_{hh}) \in [0, 0.2] \times [0, 2] \times [0, 2] \times [0, 3.8] \times \mathcal{K}^4_{n-1}, \text{ and} && (F') \\
& \mathcal{K}_n = [k^*, k^{**}_n], \text{ where } k^{**}_n := \ \max\ k : W_n(k) \geq U^2(c^*)
\end{aligned}
$$

4.A.4 Mathematica code legend

Symbol	Description
p	Probability of high state for player 2
q	Probability of high state for player 1
δ	Discount rate
σ	Relative risk aversion coefficient
G	Vector of possible utility promises
BG	All possible combinations of 4 elements from vector G
cll, clh, chl, chh	Consumption of player 1 after states are announced
CLL, CLH, CHL, CHH	Continuation utility for player 1 after states are announced
BG1, BG2, BG3, BG4	Position of each continuation utility after each state in vector G
Uaut	Autarky utility
Uauth	Autarky utility after high state
Uautl	Autarky utility after low state
Vsgsg	Vector of guesses for the values of the value function for each promise level
kBG12	Matrix with rows: $(G_i, CLL_i, CLH_i, CHL_i, CHH_i, BG1_i, BG2_i, BG3_i, BG4_i)$
VF	Value function:
ST	Constraint: $U^1(c) \geq$ k
RH1	R_H^1
RH2	R_H^2
PH1	P_H^1
PL1	P_L^1
PH2	P_H^2
PL2	R_L^2
KbgBG12	Matrix with rows $(G_i, cll_i, clh_i, chl_i, chh_i, CLL_i, CLH_i, CHL_i, CHH_i, BG1_i, BG2_i, BG3_i, BG4_i)$

4.A.5 Mathematica code

```
VF[{t_, cll_, clh_, chl_, chh_, CLL_, CLH_, CHL_, CHH_, BG1_, BG2_,
    BG3_, BG4_}] :=
  Re[q p ( u[3.8 - chh] + \[Delta] Vsgsg[[BG4]]) +
    q  (1 - p) ( u[2 - chl] + \[Delta]  Vsgsg[[BG3]]) +   (1 - q) p (
      u[2 - clh] + \[Delta]  Vsgsg[[BG2]] ) + (1 - q) (1 - p) (
      u[0.2 - cll] + \[Delta]  Vsgsg[[BG1]])];
ST[{t_, cll_, clh_, chl_, chh_, CLL_, CLH_, CHL_, CHH_, BG1_, BG2_,
    BG3_, BG4_}] :=
  q p (u[chh] + \[Delta] CHH) +
   q (1 - p) (u[chl] + \[Delta] CHL) + (1 -
      q) p (u[clh] + \[Delta] CLH ) + (1 - q) (1 -
      p) (u[cll] + \[Delta] CLL ) - t;
ST1[{t_, CLL_, CLH_, CHL_, CHH_, BG1_, BG2_, BG3_, BG4_}] :=
  ST[{t, 0.2, 2, 2, 3.8, CLL, CLH, CHL, CHH, BG1, BG2, BG3, BG4}];
RH1[{t_, cll_, clh_, chl_, chh_, CLL_, CLH_, CHL_, CHH_, BG1_, BG2_,
    BG3_, BG4_}] :=
  p (u[chh] + \[Delta] CHH) + (1 - p) (u[chl] + \[Delta] CHL) -
   p (u[clh + 1.8] + \[Delta] CLH) - (1 -
      p) (u[cll + 1.8] + \[Delta]  CLL );
RH11[{t_, CLL_, CLH_, CHL_, CHH_, BG1_, BG2_, BG3_, BG4_}] :=
  RH1[{t, 0.2, 2, 2, 3.8, CLL, CLH, CHL, CHH, BG1, BG2, BG3, BG4}];
RH2[{t_, cll_, clh_, chl_, chh_, CLL_, CLH_, CHL_, CHH_, BG1_, BG2_,
    BG3_, BG4_}] :=
  q (u[3.8 - chh] + \[Delta] Vsgsg[[BG4]]) + (1 -
      q) (u[2 - clh] + \[Delta] Vsgsg[[BG2]]) -
   q (u[2 - chl + 1.8] + \[Delta] Vsgsg[[BG3]]) - (1 -
      q) (u[0.2 - cll + 1.8] + \[Delta]  Vsgsg[[BG1]] );
RH21[{t_, CLL_, CLH_, CHL_, CHH_, BG1_, BG2_, BG3_, BG4_}] :=
  RH2[{t, 0, 0, 0, 0, CLL, CLH, CHL, CHH, BG1, BG2, BG3, BG4}];
PH1[{t_, cll_, clh_, chl_, chh_, CLL_, CLH_, CHL_, CHH_, BG1_, BG2_,
    BG3_, BG4_}] :=
  p (u[chh] + \[Delta] CHH) + (1 - p) (u[chl] + \[Delta] CHL ) - Uauth;
PH11[{t_, CLL_, CLH_, CHL_, CHH_, BG1_, BG2_, BG3_, BG4_}] :=
```

```
  PH1[{t, 0.2, 2, 2, 3.8, CLL, CLH, CHL, CHH, BG1, BG2, BG3, BG4}];
PL1[{t_, cll_, clh_, chl_, chh_, CLL_, CLH_, CHL_, CHH_, BG1_, BG2_,
    BG3_, BG4_}] :=
  p (u[clh] + \[Delta] CLH ) + (1 - p) ( u[cll] + \[Delta] CLL) -
   Uautl;
PL11[{t_, CLL_, CLH_, CHL_, CHH_, BG1_, BG2_, BG3_, BG4_}] :=
 PL1[{t, 0.2, 2, 2, 3.8, CLL, CLH, CHL, CHH, BG1, BG2, BG3, BG4}];
PL2[{t_, cll_, clh_, chl_, chh_, CLL_, CLH_, CHL_, CHH_, BG1_, BG2_,
   BG3_, BG4_}] :=
 q (u[2 - chl] + \[Delta] Vsgsg[[BG2]] ) + (1 - q) (
    u[0.2 - cll] + \[Delta] Vsgsg[[BG1]]) - Uautl;
PL22[{t_, CLL_, CLH_, CHL_, CHH_, BG1_, BG2_, BG3_, BG4_}] :=
 PL2[{t, 0, 0, 0, 0, CLL, CLH, CHL, CHH, BG1, BG2, BG3, BG4}];
PH2[{t_, cll_, clh_, chl_, chh_, CLL_, CLH_, CHL_, CHH_, BG1_, BG2_,
   BG3_, BG4_}] :=
 q (u[3.8 - chh] + \[Delta] Vsgsg[[BG4]]) + (1 -
     q) (u[2 - clh] + \[Delta] Vsgsg[[BG3]] ) - Uauth;
PH22[{t_, CLL_, CLH_, CHL_, CHH_, BG1_, BG2_, BG3_, BG4_}] :=
 PH2[{t, 0, 0, 0, 0, CLL, CLH, CHL, CHH, BG1, BG2, BG3, BG4}];
u[c_] := c^(1 - \[Sigma])/(
 1 - \[Sigma])    ; \[Delta] = 0.96 ; p = 0.5 ; \[Sigma] = 0.3 ; q = \
0.5;(* Value function and participation, incentive compatible and \
promised utility constraints*)(*Autarchy utility*)Uaut = (
   p u[1.9] + (1 - p) u[0.1]) \!\(
\*UnderoverscriptBox[\(\[Sum]\), \(t = 0\), \(\[Infinity]\)]
\*SuperscriptBox[\(\[Delta]\), \(t\)]\);(*Autarchy utility after high \
state*)Uauth = u[1.9] + (p u[1.9] + (1 - p) u[0.1]) \!\(
\*UnderoverscriptBox[\(\[Sum]\), \(t = 1\), \(\[Infinity]\)]
\*SuperscriptBox[\(\[Delta]\), \(t\)]\);(*Autarchy utility after low \
state*)Uautl = u[0.1] + (p u[1.9] + (1 - p) u[0.1]) \!\(
\*UnderoverscriptBox[\(\[Sum]\), \(t = 1\), \(\[Infinity]\)]
\*SuperscriptBox[\(\[Delta]\), \(t\)]\);(*grid length*)gl = \
50;(*difference between lowest and highest promised utility guess for \
second step*)r1 = 0.02;(*difference between lowest and highest \
promised utility initial guess*)r = 0.021;(*Before iteration promised \
utility guessed grid *)Vsgsg =
```

```
 Reverse[Table[
   k, {k, Uaut,
    Uaut*(1 + r), (Uaut*(1 + r) - Uaut)/gl}]]; ((*Iteration starts *)
 Label[begin]; Clear[kBG12c]; Clear[KbgBG12]; Clear[part];
 Clear[uruhq]; Clear[hnfg2]; Clear[tRE]; Clear[Treetye]; Clear[uruhq];
  Clear[hnfg2];(*Promised Utility values; from autarky utility to 1+r \
of autarky.  *)
 G = Table[
   k, {k, Uaut,
    Uaut*(1 + r), (Uaut*(1 + r) - Uaut)/
     gl}];(*Promised Utility permutations in sets of 4, for 4 \
possible states; matrix n x 4 *)
 kkk = Permutations[
   G, {4}];(*Reduce permuatations since some combinations are \
impossible - Select only possible combinations *)
 treeeee[i_] :=
  Select[kkk[[i]],
   kkk[[i, 1]] < kkk[[i, 3]] && kkk[[i, 2]] < kkk[[i, 4]] &&
     kkk[[i, 2]] < kkk[[i, 1]] &&
     kkk[[i, 4]] <
      kkk[[i, 3]] &];(*Create a matrix with all the possible \
permutations of promised utilities  *)
 tt = Table[
   treeeee[i], {i, 1,
    Length[kkk]}];(*Remove the empty rows of the matrix above
*)
 BG = Select[tt,
   UnsameQ[#, {}] &];(*function that generates vectors of same level \
of promised utilities repeated length BG, for all levels of promised \
utilities*)
 g11[i_] :=
  ConstantArray[G[[i]],
   Length[BG]];(*Create a column vector of all the elements generated \
by the function above, Length[BG}x1*)
 kvaluesconstr =
  Partition[Flatten[Table[g11[i], {i, Length[G]}]],
```

```
   1];(*Replicate BG Length[G] times and combine into matrix \
dimensions Length[BG]*Length[G]x4*)
 TRet = Partition[Flatten[Table[BG, {Length[G]}]],
   4];(*Create matrix length[Tret]x5, where the first column is G \
repeated BG times and the rest is the matrix cretaed above*)
 kBG = Join[kvaluesconstr, TRet,
   2];(*Create matrix that shows the corresponding position of \
elements in BG in vector G. dimenstions Length[BG]x4*)
 yiu = Partition[Flatten[Position[G, #] & /@ Flatten[BG]],
   4];(*Replicate the matrix above Length[G] times and cosntruct \
matrix Length[BG]*Length[G]x4   *)
 Putr = Partition[Flatten[Table[yiu, {Length[G]}]],
   4];(*Join the last two created matrices into one matrix \
Length[BG]*Length[G]x9. example of row i {Subscript[G, \
i],Subscript[BG, i1,] Subscript[BG, i2,] Subscript[BG, \
i3,] Subscript[BG, i4,] Position[G,Subscript[BG, \
i1]],Position[G,Subscript[BG, i2]]...}*)kBG12 = Join[kBG, Putr, 2];
 Clear[kkk]; Clear[tt]; Clear[kBG ]; Clear[yiu]; Clear[TRet];
 Clear[kvaluesconstr]; Clear[treeeee]; Clear[Putr];
 Clear[g11]; << Developer`;
 kBG12 = Developer`ToPackedArray[kBG12,
   Real]; kBG12;(*From the matrix built above select only the rows \
that weakly satisfy all lower thresholds of the constraints. *)
 kBG12c = Select[kBG12, ST1[#] >= 0 &];
 kBG12c = Select[kBG12, RH11[#] >= 0 &];
 kBG12c = Select[kBG12, PH11[#] >= 0 &];
 kBG12c = Select[kBG12, PL11[#] >= 0 &];
 kBG12c = Select[kBG12, RH21[#] >= 0 &];
 Clear[kBG12];(*Find the argmax of static maximization problem \
assuming promises as in the matrix above and assuming that \
continuation value for maximizer as the guesed utility above (for the \
first iteration this is the inverse of G. That is if promice CLH is \
given, this corresponds to a position in G, which also corresponds to \
the same position in the value function.) *)
 eeqnw[{t_, CLL_, CLH_, CHL_, CHH_, BG1_, BG2_, BG3_,
    BG4_}] :=   {cll, clh, chl, chh} /.
```

```
Quiet[FindMaximum[{Re[
      q p ( u[3.8 - chh] + \[Delta]  Vsgsg[[BG4]]) +
       q  (1 - p) (
         u[2 - chl] + \[Delta]  Vsgsg[[BG3]]) +   (1 - q) p (
         u[2 - clh] + \[Delta]  Vsgsg[[BG2]]) + (1 - q) (1 - p) (
         u[0.2 - cll] + \[Delta]  Vsgsg[[BG1]])],
     q p (u[chh] + \[Delta] CHH) +
        q (1 - p) (u[chl] + \[Delta] CHL) + (1 -
           q) p (u[clh] + \[Delta] CLH ) + (1 - q) (1 -
           p) (u[cll] + \[Delta] CLL ) - t >= 0 &&
      p (u[chh] + \[Delta] CHH) + (1 - p) (u[chl] + \[Delta] CHL) -
        p (u[clh + 1.8] + \[Delta] CLH) - (1 -
           p) (u[cll + 1.8] + \[Delta]  CLL ) >= 0 &&
      q (u[3.8 - chh] + \[Delta] Vsgsg[[BG4]]) + (1 -
           q) (u[2 - clh] + \[Delta] Vsgsg[[BG2]]) -
        q (u[2 - chl + 1.8] + \[Delta] Vsgsg[[BG3]]) - (1 -
           q) (u[0.2 - cll + 1.8] + \[Delta]  Vsgsg[[BG1]] ) >= 0 &&
      p (u[clh] + \[Delta] CLH ) + (1 - p) (
          u[cll] + \[Delta] CLL) - Uautl >= 0 && cll > 0 &&
      cll < 0.2 && clh > 0 && clh < 2 && chl > 0 && chl < 2 &&
      chh > 0 && chh < 3.8}, {cll, clh, chl, chh}]][[2]];
(*Apply the function to every row of the matrix. This generates \
argmaxizers for each state. *)
trhv = ParallelMap[eeqnw,
   kBG12c];(*Replace all empty outputs with 4 0s*)
trhv = trhv /. {} -> {0, 0, 0, 0};
(*Add the outputs from the equation before to the matrix with \
promices*)tewtw =
  Join[trhv, kBG12c,
   2];(*Rearrange the matrix above so that the argmaxs are placed in \
position 2,3,4,5. (This is so that the function inputs are correct \
for the functions as defined.)*)
tewtw2 =
  Drop[Transpose[Insert[Transpose[tewtw], tewtw[[All, 5]], 1]],
   0, {6}];(*Change all the rows with some zeros so that all the \
values are zero in these rows. Dimensions Length[BG]*Length[G] x 13 *)
```

```
 KbgBG12 =
  tewtw2 /. {_, 0, 0, 0, 0, _, _, _, _, _, _, _, _} -> {0, 0, 0, 0, 0,
     0, 0, 0, 0, 0, 0, 0, 0}; Clear[ifff]; Clear[trhv]; Clear[tewtw];
  Clear[eeqnw];
 Clear[tewtw2]; KbgBG12;(*Partition the matrix above  into matrices \
with Length[BG] rows. Hence, each partition corresponds to all \
possible continuations given a level of promise. *)
 part[i_] :=
  Partition[KbgBG12, Length[BG]][[
   i]];(* For each partition above select all rows that weakly \
satisfy the constraints*)
 uruhq[i_] :=
  Select[part[i],
   ST[#] >= 0 && RH1[#] >= 0  && PH1[#] >= 0 && PL1[#] >= 0 &&
     RH2[#] >=
      0 & ];(*Create a list that contains combines all the partitions \
above, dimenstions Length[G]x n x 13*)hnfg2 = Array[uruhq, Length[G]];
  VF[{}] :=
  0;(* Create a function that evaluates VF for all rows of hnfg2 \
above. Then choose the maximal value for each of the partitions made \
before. That  is for each promise Subscript[G, i] the maximum VF \
given grid.*)
 tRE[i_] :=
  Max[ParallelMap[VF, hnfg2, {2}][[
    i]]];(* create a vector with these values, length G NOTE: I \
CHANGED THE CODE BELOW BY FROM := T0 = ON 28/8/20.   *)
 Treetye =
  Table [tRE[i], {i, 1,
    Length[G]}];(* Calculate growth of largest from smallest values \
in the vector above *)
 r1 = -1 + Treetye [[1]]/
   Treetye[[
    Length[Vsgsg]]];(* find the line (intercept and slope) that \
satisfies both growth rates, the r from initial value function guess \
```

```
and the r found before from new value function guess.*)
 sol = Flatten[{\[Alpha], \[Beta]} /.
    Solve[Treetye[[Length[G]]] == \[Alpha] + \[Beta]*r &&
      Vsgsg[[Length[G]]] == \[Alpha] + \[Beta]*
         r1, {\[Alpha], \[Beta]}]]; a = sol[[1]]; b = sol[[2]];
 (* r1 becomes the initial r that was used to generate in this round \
G *)r1 = r;
 Clear[r];(* Given a and b found above r is the value required to get \
G[[1]]. This is the next iteration's r. Note that as the iterations \
progress the vectors Vsgsg and Treetye change from initial values.
*)
 rf = Solve[G[[1]] == a + b*r, r][[1]];
 r = r /. rf;(*As long as the maximum absolute distance between Vsgsg \
and Treetye is more than 10^(-3) replace Vsgsg with Treetye, then \
print r and Vsgsg *)Print[Vsgsg];
 While[Max[Abs[Treetye - Vsgsg]] > 0.0005,
  Vsgsg = Replace[Vsgsg, Vsgsg -> Treetye]; Print[r];
  Print[Treetye]; Goto[begin]];
 Print[r];
 Print[Treetye];(*Finally if the difference between the lowest value \
of the guessed value function grid and the autarchy utility is \
greater than 0.005 do another iteration, otherwise stop.*)
 While[Abs[Vsgsg[[Length[G]]] - G[[1]]] > 0.0003, Goto[begin]])
```

References

Abdulkadiroğlu, Atila and Kyle Bagwell (2013). "Trust, reciprocity, and favors in cooperative relationships". In: *American economic journal: Microeconomics* 5.2, pp. 213–59.

Ábrahám, Árpád and Sarolta Laczó (2018). "Efficient risk sharing with limited commitment and storage". In: *The Review of Economic Studies* 85.3, pp. 1389–1424.

Acharya, Viral V (2009). "A Theory of Systemic Risk and Design of". In: *Journal of Financial Stability* 5.3, pp. 224–255. URL: http://econpapers.repec.org/RePEc:eee:finsta:v:5:y:2009:i:3:p:224-255.

Acharya, Viral V, Hamid Mehran, and Anjan V Thakor (2016). "Caught between Scylla and Charybdis? Regulating bank leverage when there is rent seeking and risk shifting". In: *The Review of Corporate Finance Studies* 5.1, pp. 36–75.

Acharya, Viral V. and Tanju Yorulmazer (2007). "Too many to fail-An analysis of time-inconsistency in bank closure policies". In: *Journal of Financial Intermediation* 16.1, pp. 1–31. ISSN: 10429573. DOI: 10.1016/j.jfi.2006.06.001.

Allen, Franklin, Ana Babus, and Elena Carletti (2012). "Asset commonality, debt maturity and systemic risk". In: *Journal of Financial Economics* 104.3, pp. 519–534. ISSN: 0304405X. DOI: 10.1016/j.jfineco.2011.07.003.

Allen, Franklin and Douglas Gale (2000). "Financial contagion". In: *Journal of political economy* 108.1, pp. 1–33.

Alonso, Ricardo and Niko Matouschek (2008). "Optimal delegation". In: *Review of Economic Studies* 75.1, pp. 259–293. ISSN: 00346527. DOI: 10.1111/j.1467-937X.2007.00471.x.

Argenziano, Rossella, Sergei Severinov, and Francesco Squintani (2016). "Strategic Information Acquisition and Transmission". In: 8.3, pp. 119–155.

Brocas, Isabelle and Juan D. Carrillo (2007). "Influence through ignorance". In: *RAND Journal of Economics* 38.4, pp. 931–947. ISSN: 07416261. DOI: 10.1111/j.0741-6261.2007.00119.x.

Cabrales, Antonio, Piero Gottardi, and Fernando Vega-Redondo (2017). “Risk sharing and contagion in networks”. In: *Review of Financial Studies* 30.9, pp. 3086–3127. ISSN: 14657368. DOI: 10.1093/rfs/hhx077.

Che, Yeon-Koo and Konrad Mierendorff (2019). “Optimal dynamic allocation of attention”. In: *American Economic Review* 109.8, pp. 2993–3029.

Dasgupta, Amil (2004). “Financial contagion through capital connections: A model of the origin and spread of bank panics”. In: *Journal of the European Economic Association* 2.6, pp. 1049–1084. ISSN: 15424766. DOI: 10.1162/1542476042813896.

Deimen, Inga and Dezsö Szalay (2015). “Information, authority, and smooth communication in organizations”. In.

Dell, Giovanni and Lev Ratnovski (2012). “Bailouts , Contagion , and Moral Hazard”. In: pp. 1–21.

Dessein, Wouter (2002). “Authority and Communication in Organizations”. In: *Review of Economic Studies* 69, pp. 811–838.

Diamond, Douglas W and Philip H Dybvig (1983). “Bank runs, deposit insurance, and liquidity”. In: *Journal of political economy* 91.3, pp. 401–419.

Erol, Selman (2017). “Network Hazard and Bailouts”. In: *SSRN Electronic Journal*. DOI: 10.2139/ssrn.3034406.

Escobar, Juan F and Qiaoxi Zhang (2021). “Delegating learning”. In: *Theoretical Economics* 16.2, pp. 571–603.

Farhi, Emmanuel and Jean Tirole (2012). “Collective moral hazard, maturity mismatch, and systemic bailouts”. In: *American Economic Review* 102.1, pp. 60–93. ISSN: 00028282. DOI: 10.1257/aer.102.1.60.

Freixas, Xavier, Bruno M. Parigi, and Jean-Charles Rochet (2000). “Systemic Risk, Interbank Relations, and Liquidity Provision by the Central Bank”. In: *Journal of Money, Credit and Banking* 32.3, p. 611. ISSN: 00222879. DOI: 10.2307/2601198.

Garfagnini, Umberto (2011). “Delegated Experimentation”. In: *Unpublished Manuscript*, pp. 1–46. URL: http://www.kellogg.northwestern.edu/faculty/garfagnini/Garfagnini{_}JMP.pdf.

Guo, Yingni (2016). “Dynamic delegation of experimentation”. In: *American Economic Review* 106.8, pp. 1969–2008. ISSN: 00028282. DOI: 10.1257/aer.20141215.

Herresthal, Claudia (2022). “Hidden testing and selective disclosure of evidence”. In: *Journal of Economic Theory*, p. 105402.

Hertel, Johanna (2004). “Efficient and sustainable risk sharing with adverse selection”. In: *Princeton University*.

Ke, Tony and J Miguel Villas-Boas (2017). “Optimal learning before choice”. In.

Keller, Godfrey, Sven Rady, and Martin Cripps (2005). "Strategic experimentation with exponential bandits". In: *Econometrica* 73.1, pp. 39–68.

Kocherlakota, Narayana R (1996). "Implications of efficient risk sharing without commitment". In: *The Review of Economic Studies* 63.4, pp. 595–609.

Koeppl, Thorsten V (2007). "Optimal dynamic risk sharing when enforcement is a decision variable". In: *Journal of Economic Theory* 134.1, pp. 34–60.

Ligon, Ethan, Jonathan P Thomas, and Tim Worrall (2002). "Informal insurance arrangements with limited commitment: Theory and evidence from village economies". In: *The Review of Economic Studies* 69.1, pp. 209–244.

Mayskaya, Tatiana (2020). "Dynamic choice of information sources". In: *California Institute of Technology Social Science Working Paper, ICEF Working Paper WP9/2019/05.*

Möbius, Markus (2001). "Trading favors". In: *Unpublished paper.[283].*

Nikandrova, Arina and Romans Pancs (2018). "Dynamic project selection". In: *Theoretical Economics* 13.1, pp. 115–143.

Nikandrova, Arina and Jevgenijs Steinbuks (2017). "Contracting for the second best in dysfunctional electricity markets". In: *Journal of Regulatory Economics* 51.1, pp. 41–71.

Thomas, Jonathan and Tim Worrall (1988). "Self-enforcing wage contracts". In: *The Review of Economic Studies* 55.4, pp. 541–554.

Wald, Abraham (1947). "Foundations of a general theory of sequential decision functions". In: *Econometrica, Journal of the Econometric Society*, pp. 279–313.

Zhong, Weijie (2022). "Optimal dynamic information acquisition". In: *Econometrica* 90.4, pp. 1537–1582.